CANADIAN THESES ON MICROFICHE

I.S.B.N.

THESES CANADIENNES SUR MICROFICHE

National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada
Direction du développement des collections

Service des thèses canadiennes
sur microfiche

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

NL-339 (r. 82/08)

Canada

0-315-19396-4

National Library of Canada
Canadian Theses Division
Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada
Division des thèses canadiennes

48

67304

PERMISSION TO MICROFILM — AUTORISATION DE MICROFILMER

• Please print or type — Écrire en lettres moulées ou dactylographier

Full Name of Author — Nom complet de l'auteur
Nobuya ITAGAKI

Date of Birth — Date de naissance
Nov. 1st, 1951

Country of Birth — Lieu de naissance
Japan

Permanent Address — Résidence fixe
173, Natori, Murayama-shi; Yamagata-Ken Japan 995

Title of Thesis — Titre de la thèse
Inherent nominal properties as determinants of grammatical subjects.

University — Université
University of Alberta

Degree for which thesis was presented — Grade pour lequel cette thèse fut présentée
Ph. D.

Year this degree conferred — Année d'obtention de ce grade
1984

Name of Supervisor — Nom du directeur de thèse
Dr. Gary D. Prideaux

Permission is hereby granted to the NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈQUE NATIONALE DU CANADA de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans l'autorisation écrite de l'auteur.

Date
Sept. 12 1984

Signature
N. Itagaki

NL-91 (4/77)

THE UNIVERSITY OF ALBERTA

INHERENT NOMINAL PROPERTIES AS DETERMINANTS OF GRAMMATICAL SUBJECTS

by

©

NOBUYA ITAGAKI

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY
IN
PSYCHOLINGUISTICS

DEPARTMENT OF LINGUISTICS

EDMONTON, ALBERTA
FALL, 1984

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: NOBUYA ITAGAKI

TITLE OF THESIS: INHERENT NOMINAL PROPERTIES AS DETERMINANTS OF GRAMMATICAL SUBJECTS

DEGREE FOR WHICH THESIS WAS PRESENTED: DOCTOR OF PHILOSOPHY

YEAR THIS DEGREE GRANTED: FALL, 1984

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(Signed) N. Itagaki

PERMANENT ADDRESS:

173, Natori, Muraya-Shi
Yamagata-Ken, JAPAN, 995

DATED August 13, 1984

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled INHERENT NOMINAL PROPERTIES AS DETERMINANTS OF GRAMMATICAL SUBJECTS submitted by NOBUYA ITAGAKI in partial fulfilment of the requirements of the degree of DOCTOR OF PHILOSOPHY in PSYCHOLINGUISTICS.

Gary Prideaux
Supervisor

John T. Hogan

Matthew S. Dryer

Joseph F. Kess
External Examiner

Date August 13, 1984

For the people who taught me there is always hope in life...

ABSTRACT

This thesis investigates the following non-discourse properties of nominal referents as important determinants of English grammatical subjects from a psycholinguistic point of view: frequency, concreteness, animacy, and prototypicality. The basic hypothesis of this work is that the more frequent, concrete, animate, and/or prototypical a nominal referents, the more likely it is to be chosen as a grammatical subject. This hypothesis was tested empirically, using an experimental paradigm of sentence construction.

It was found that there were significant interactions among frequency, concreteness and animacy with regard to the choice of grammatical subjects. Specifically, animacy and concreteness were both found to be important non-discourse determinants of subject selection with respect to high frequency nominal referents, but not with respect to low frequency ones. Moreover, frequency turned out to be an important determinant only of human nominal referents. Prototypicality was shown to be another major determinant, but it has a smaller effect than was expected.

The psycholinguistic implications of these experimental findings are discussed. First, the notion of grammatical subject is further clarified. Second, a psychological explanation of a nearly universal relative order of subjects before other nominal constituents is offered.

ACKNOWLEDGMENTS

I would like to express my deep gratitude to my thesis supervior, Dr. Gary D. Prideaux for his patient guidance, and his encouragement throughout my five years of graduate study at the Department of Linguistics, University of Alberta.

I would like to express my sincere gratitude to the other committee members: Dr. M.S. Dryer, Dr. K.T. Holden, Dr. J.F. Kess (of the University of Victoria), and Dr. L. Marckworth Stanford for valuable suggestions in completing this thesis. I am particularly indebted to Dr. John T. Hogan for his advice on the experimental design and statistical analyses of this study.

My deep appreciation is extented to Mr. P.F. Assmann, Ms. T.M. Derwing, Mr. M.J. Munro, and other fellow graduate students who never hesitated to offer assistance to me throughout this thesis. (I hope that they will finish their own work soon.)

I thank my friends and teachers in Japan for all the support they have shown to me. Finally, I thank the Department of Linguistics for financially supporting my stay and graduate study here in Edmonton.

TABLE OF CONTENTS

CHAPTER PAGE

1. INTRODUCTION 1
2. REVIEWS OF RELATED STUDIES 18
  2.1 The Cognitive Availability of Nominal Referents 19
  2.2 Grammatical Subject as Cognitive Point of Departure 39
    2.2.1 Linguistic Studies of Cognitive Point of Departure 41
    2.2.2 Psycholinguistic Studies of Cognitive Point of Departure 56
  2.3 Some Implications for Linguistics 71
3. EXPERIMENTS 82
  3.1 Experiment 1 83
    3.1.1 Purpose 83
    3.1.2 Method 83
    3.1.3 Analyses and Results 97
    3.1.4 Discussion and Conclusions 107
  3.2 Experiment 2 112
    3.2.1 Purpose 112
    3.2.2 Method 112
    3.2.3 Analyses and Results 117
    3.2.4 Discussion and Conclusions 123
  3.3 Experiment 3 125

3.3.1 Purpose 125
3.3.2 Method 125
3.3.3 Analyses and Results 126
3.3.4 Discussion and Conclusions 130
4. General Discussion and Conclusions 132
4.1 Sumary and Conclusions of Results 132
4.2 Psycholinguistic Contributions
to Linguistics 147
4.3 Some Limitations of the Present Thesis 151

REFERENCES 154
APPENDIX A. NOUNS USED IN EXPERIMENT 1 166
APPENDIX B. SAMPLE SINGLE SENTENCES 168
APPENDIX C. SAMPLE PASSAGES 171
APPENDIX D. MEAN RATIOS FOR NOUNS 173
APPENDIX E. DETAILED RESULTS OF *A POSTERIORI*
TESTS 175
APPENDIX F. PICTURES USED IN EXPERIMENTS 2 and 3 176

LIST OF TABLES

Table	Description	Page
1.	Mean Ratios and Standard Deviations	91
2.	Results of Analysis of Variance (Ratios)	93
3.	Results of *A Posteriori* Tests (Ratios)	95
4.	Results of *A Posteriori* Tests (Ratios)	95
5.	Mean Percentages and Standard Deviations	100
6.	Results of Analysis of Variance (Percentages)	101
7.	Results of *A Posteriori* Tests (Percentages)	105
8.	Results of *A Posteriori* Tests (Percentages)	105
9.	Predictions of the Three Factors and Mean Markedness Scores and Standard Deviations (Experiment 2)	122
10.	Results of *Sign* Tests on Six Pairs (Experiment 2)	122
11.	Predictions of the Three Factors and Mean Markedness Scores and Standard Deviations (Experiment 3)	128
12.	Results of *Sign* tests on Six Pairs (Experiment 3)	128

LIST OF FIGURES

Figure	Description	Page
1.	Interaction between Animacy-Concreteness and Frequency	94
2.	Main Effect of Task	102
3.	Interaction between Animacy-Concreteness and Frequency	103

1. INTRODUCTION

Why particular noun phrases are chosen as grammatical subjects in sentences is the question to be addressed in this thesis. Some typical examples of alternative English subjects are:

1) a. The shining wampum attracted the woman.
   b. The woman was attracted by the shining wampum.
2) a. The greatest human invention is the locomotive.
   b. The locomotive is the greatest human invention.
3) a. The Phillies beat the Dodgers.
   b. The Dodgers lost to the Phillies.

Both sentences of each pair refer to the same *state of affairs*, but they have different grammatical subjects. It may be argued that (3 a,b) differ in the propositional content as well, namely they have different agents, *The Phillies* and *The Dodgers*, respectively. The present thesis does not focus on the choice of propositional content, but rather on the choice of grammatical subjects (cf. Itagaki, 1982).

The question under consideration here is how a speaker chooses between each of the three pairs of noun phrases in the sample sentences above: *the shining wampum* and *the woman*, *the greatest human invention* and *the locomotive*, and *the Phillies* and *the Dodgers*. The primary purpose of this work is to address this question from a psycholinguistic point of view. The following specific issues will be examined in this thesis: (a) What *psychological function* does a grammatical subject serve? (b) Which *psychological factors* are reflected in a speaker's choice of grammatical subjects?

The present work is also aimed at articulating the psycholinguistic implications associated with certain linguistic phenomena. First, it may be possible to contribute to the clarification of the notion of grammatical subject, specifically in conjunction with Keenan's (1976) work on the "universal definition" of subject. Second, it may be possible to offer a partial psycholinguistic account for a nearly universal relative order of grammatical subjects before other nominal constituents, for example objects (e.g., Greenberg, 1966).

Traditionally, grammatical subjects have been discussed in terms of the so-called "cognitive point of departure," "starting point," or "psychological subject" in the mind or consciousness of a speaker (Firbas, 1974; Gundel, 1977; Lyons, 1977; Mathesius, 1964; Mathews, 1981; Sandmann, 1979). More specifically, the speaker's cognitive point of departure tends to be realized by the grammatical subject rather than by other grammatical roles in a sentence. The notion of the speaker's cognitive point of departure was originally proposed by several linguists and psychologists of the 19th century, such as von der Gadelentz, Herman Paul, and Henri Weil. These scholars generally assumed that the flow of ideas in the mind of the speaker might start with some particular idea or set of ideas which they called the speaker's cognitive point of departure.

It has also been assumed that the speaker's cognitive point of departure refers to the idea or set of ideas that is *most available* in the mind (or consciousness) of the speaker (Firbas, 1974; Krupa, 1982; Lyons, 1977). Firbas, for example, explains Weil's notion of point of departure as follows:

> A sentence contains a point of departure (an initial notion) and a goal of discourse. The point of departure is equally present to the speaker and the hearer; it is their rallying point, the ground on which they meet. The goal of discourse presents the very information that is to be imparted to the hearer. Weil claims that the movement from the initial notion to the goal of discourse reveals the movement of the mind itself. (p. 12)

Furthermore, Lyons (1977) states: "... the cognitive point of departure - the entity or topic that the speaker had in mind when he formulated the intention to produce an utterance" (p. 508). From these comments, it can be surmised that the cognitive point of departure is some kind of psychological element in the minds of the speaker and his hearers from which they may proceed to formulate and comprehend the sentence, respectively. Moreover, the cognitive point of departure as such is the one most available or most established in the minds of the speaker and his hearers (Ertel, 1977; Keenan, 1978; Krupa, 1982; MacWhinney, 1977).

Following these considerations, this thesis makes two basic assumptions. First, a primary psychological function of a grammatical subject is to refer to a speaker's cognitive point of departure, which can be assumed to be most available in the mind of the speaker. Second, the more likely a nominal referent/concept is to be chosen by a speaker as his cognitive point of departure, the more likely it will be chosen by him as a grammatical subject[1].

It may be helpful for the understanding of these points to return to the examples (1-3 a,b). Suppose that *for some reason*, the referents of *the woman, the locomotive* and *the Dodgers* are more available in the mind of a speaker than those of *the shining wampum, the greatest human invention* and *the Phillies*. It would then be likely for the speaker to choose the former as the cognitive points of departure than the latter. As a result, he might utter (1b), (2b), and (3b) rather than (1a), (2a), and (3a), respectively.

A crucial question must be raised at this point: Why are some particular nominal referents more available in the mind of a speaker? Put another way, how does a speaker choose certain nominal referents as his cognitive point of departure rather than others, with the former being realized as grammatical subjects? Returning to the above examples, why are the referents of *the woman, the locomotive* and *the Dodgers* more available in the mind of the speaker than those of *the shining wampum, the greatest human invention* and *the Phillies*, respectively? There is no doubt that the answers to these questions will contribute significantly to a psycholinguistic account of the speaker's choice of grammatical subjects.

---

[1] Throughout this thesis, the terms "nominal referent" and "nominal concept" will be used following Ogden and Richards (1923).

One possible answer is that the speaker's cognitive point of departure is contextually or situationally established as part of the "given information" as opposed to the "new information" (e.g., Allerton, 1978, 1979; Chafe, 1974, 1976; Clark & Clark, 1977, 1978; Givón, 1979a, 1979b, 1983; Keenan, 1976; Tomlin, 1979). It is generally assumed by these authors that given information is shared with the hearers. It seems plausible, therefore, for the speaker to choose given information as his cognitive point of departure rather than new information, since given information is, for the most part, most available in the minds of both the speaker and the hearer (Clark & Clark, 1977; Grice, 1975).

If this is so, it can be hypothesized that grammatical subjects are likely to refer to given information rather than new information. This hypothesis has indeed been supported in a large number of psycholinguistic studies (e.g., Bock, 1977; Bock & Irwin, 1980; Clark & Haviland, 1974; Engelkamp & Zimmer, 1983; Hornby, 1974; Osgood, 1971; Osgood & Bock, 1977). However, this line of psychological accounting for grammatical subjects is not the primary concern of the present work. Instead, this work is focused on a second aspect of the psycholinguistic notion of grammatical subject.

This second aspect of a psychological account for a speaker's cognitive point of departure has to do with the assumption that some nominal referents/concepts are easier to process (i.e., learn, recall, recognize, and so on) at a cognitive level than others (Bock, 1982; Brown, 1958; Kempen, 1977, 1978; Kempen & Huijbers, 1983; Levelt & Hassan, 1981). It has actually been shown that people perform better with some verbal or nonverbal stimuli (e.g., words and pictures) than with others in various cognitive tasks, such as paired-associate learning, recall, recognition, word production, and the like (e.g., Brown, 1976; Brown & Lenneberg, 1954; Gregg, 1976; Paivio, 1969, 1971; Paivio & Begg, 1981; Rosch, 1978; Underwood & Schulz, 1960).

Following these experimental findings, it seems reasonable to assume that nominal referents/concepts themselves differ in their *availability* for human cognition in the sense that they represent differing degrees of difficulty or ease of cognitive processing (Bock, 1982;

Brown, 1958, 1976; Brown & Lenneberg, 1954). From the viewpoint of sentence production, this can be taken to mean that a speaker may have more *cognitive access* to some nominal referents/concepts than others in the sense that the former are easier to process at a cognitive level than are the latter.

Brown and Lenneberg (1954), for example, introduce the notion of cognitive availability as follows:

> It would mean - to cite another example - that the Hopi is less often called upon to distinguish airplanes, aviators, and butterflies than is the American, since the Hopi has but a single name for all three of these. Such conclusions are, of course, supported by extralinguistic cultural analysis, which reveals the importance of snow in the Eskimo's life and the comparative indifference of the Hopi to airplanes and aviators.
>
> We will go further and propose that *increased frequency of a perceptual categorization will mean a generally greater "availability" of that category*. In the experimental study of memory, we are accustomed to think of the methods of recall, recognition, and relearning as increasingly sensitive indices of retention. (emphasis added, pp. 455-456)

It should be noted that Brown and Lenneberg's notion of cognitive availability is the primary basis for the assumption made in this thesis, except that they define the cognitive availability of nominal referents only in terms of their frequencies of occurrence. That is, the more frequently people come across certain nominal referents in their everyday life, the more cognitive access they may have to those referents.

This work assumes that there may be a certain positive correlation between the cognitive availability of nominal referents and their likelihood of being cognitive points of departure in the minds of speakers. The primary assumption is then restated as follows: some nominal referents are easier to process and thus are more available in the mind of a speaker than are others. That is, the speaker tends to choose nominal referents of high cognitive availability as his cognitive points of departure rather than those of low cognitive availability.

We have so far discussed two accountings for the cognitive point of departure in the mind of a speaker as a primary psychological function of a grammatical subject. According to the first, it is the discourse property, for example given information, topic, and so on, of nominal referents that determines the cognitive point of departure in the minds of speakers. The second claims that a speaker bases his choice of cognitive point of departure on the non-discourse aspect (i.e., cognitive availability) of nominal referents. That is, there are basically two different types of determinants of the cognitive points of departure in the minds of speakers: "discourse" (see Itagaki, 1982) and "non-discourse" determinants, the second of which is the primary concern here.

The basic hypothesis of the present work is as follows: the higher the cognitive availability of nominal referents/concepts, the more likely they are to be chosen by speakers as their cognitive points of departure and thus as grammatical subjects. This hypothesis will be experimentally tested particularly with regard to English sentences.

Although discourse determinants (i.e., discourse relations) have been extensively studied in both the linguistic and psycholinguistic literatures (e.g., Firbas, 1974; Givón, 1983; Halliday, 1970; Clark & Haviland, 1977; Clark & Clark, 1977), non-discourse determinants (i.e., cognitive availability) have not been explored. This work will investigate the non-discourse determinants of the cognitive point of departure in the mind of a speaker and thus of the grammatical subject (cf. Itagaki, 1982; Tomlin, 1983).

It is next necessary to elaborate further the notion of cognitive availability (i.e., non-discourse determinants) of nominal referents. Bock (1982), for example, identifies the following two major determinants of the lexicalization availability of nominal referents: "conceptual (or semantic) accessibility" and "phonological accessibility."

The phonological accessibility of nominal concepts may be related, for example, to the numbers of syllables in the corresponding words. That is, words with smaller numbers of syllables may be easier to retrieve and produce than those with larger numbers of syllables (Zipf, 1935). In this sense, a speaker may have more lexical access to the former than the latter.

On the other hand, the conceptual (or semantic) accessibility of nominal referents has to do with their *inherent properties* which may contribute to their ease (or difficulty) of cognitive processing. In this connection, Bock (1982) states:

> One variable with a strong claim to a relationship to ease of lexicalization is concreteness. *Concrete objects and events are certainly better coded (in the sense of Rosch et al., 1976) in the standard lexicons of the languages of the world than the apparently more open set of possible abstraction*, perhaps because the domains of abstract categories lack the correlational structure that underlies the formation of basic-level concepts... The general difficulty created by the need to talk about abstract versus concrete topics in sentence production has been demonstrated in a number of experiments... (emphasis added, p. 17)

Furthermore, it is suggested that the notion of conceptual accessibility is to be further sorted out with regard to other possible variables.

There is no doubt that there are other possible inherent properties of nominal referents that may be related to their availability for human cognition, specifically as the non-discourse determinants of availability in the mind of a speaker. The present thesis is concerned with the cognitive availability of nominal referents as defined in terms of four inherent non-discourse properties: *frequency*, *concreteness*, *animacy*, and *prototypicality*. These four properties seem to be reasonably representative of the availability of nominal referents for human cognition in the sense that it has been suggested that one has less difficulty in processing more frequent, concrete, animate, and/or prototypical nominal referents at a cognitive level than less frequent, concrete, animate, and/or prototypical nominal ones, respectively (e.g., Bock, 1982; Brown & Lenneberg, 1954; Paivio, 1971; Paivio & Begg, 1981; Rosch, 1973, 1978). It seems reasonable, in turn, to hypothesize that the four inherent properties of nominal referents are major parameters of their likelihood of being speakers' cognitive points of departure and thus grammatical subjects.

It must be mentioned that these four nominal properties might not all be considered to be "inherent" in referents themselves in the same sense. It might be suggested that frequency and prototypicality are not inherent properties of referents in the same sense that animacy and concreteness are. That is, while animacy and concreteness can be considered as a part of the denotation of words, frequency is a performance phenomenon referring to usage. Similarly, prototypicality has to do with speakers' judgments made among alternative members of a given family or set of words.

However, througout this thesis, frequency, concreteness, animacy and prototypicality will be referred to as inherent nominal properties in the sense that they are not related to the discourse, but to the non-discourse aspects of nominal referents. This thesis is not concerned with whether or not each of the four properties is inherent in nominal referents, but whether each of them is a discourse or non-discourse property. This point will be further discussed in Section 2.1.

It is not unreasonable to assume that people have acquired more knowledge about frequent objects and things than infrequent objects and things, since the more frequently they encounter some nominal referents in the external and internal world, the more likely it is that they understand and know them well. Consequently, it is plausible that they could imagine, recognize, and remember frequent objects and things much better than infrequent objects and things as a function of the so-called "frequency effect" (Gregg, 1976; Oldfield, 1963, 1965). It is on the basis of this line of reasoning that the frequency of nominal referents is hypothesized as a parameter of cognitive availability in the mind of a speaker.

As stated earlier, Brown (1958, 1976) and Brown and Lenneberg (1954) define the cognitive availability of nominal referents only in terms of their frequency of occurrence, i.e., as a function of frequency only. However, it will be clear below that such a narrow view of cognitive availability is hard to justify.

It is difficult to have direct nonverbal experience with abstract nominal referents, such as justice, virtue, honesty, and so on, in the external world. By contrast, people can have both direct verbal and nonverbal experience with concrete nominal referents in the external and internal world. Therefore, it seems reasonable to assume that people may have acquired more knowledge about concrete nominal referents than about nominal referents. It has been reported that people can process concrete noun stimuli much more easily at a cognitive level than abstract ones (Kieras, 1978; Paivio & Begg, 1981; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). Accordingly, the concreteness of nominal referents can be hypothesized as a second parameter of their availability for human cognition, particularly in the mind of a speaker.

Traditionally, it has been argued that people are motivated to perceive and think about their external and internal world primarily from their own point of view, namely in terms of the so-called "egocentrism" and "anthropocentrism" in human thinking. People, for example, seem to be more concerned with humans and human-related entities than inanimate entities. That is, their normal, daily verbal and nonverbal experience is likely to be centered on animate entities rather than inanimate ones. It, therefore, seems reasonable to hypothesize the animacy of nominal referents as a third parameter of a speaker's cognitive access (Hörmann, 1981; Piaget, 1926; Zubin, 1979).

The present thesis is also concerned with the prototypicality of nominal referents as a fourth parameter of their cognitive availability. People may, for example, agree that cars are more typical of vehicles than trailers or carriages and that robins and sparrows are more typically birds than are chickens and ducks. This means that cars and robins are more consistent with our opinions or images about what vehicles and birds, respectively, are like than the others. More importantly, it has been shown that people can process or perform with the central members (e.g., cars and robins) of natural basic categories more easily in cognitive tasks than peripheral members of the categories (e.g., carriages and ducks) (Battig & Montague, 1969; Rosch, 1973, 1974, 1975, 1978; Rosch et al., 1976). Accordingly, it seems conceivable to hypothesize the prototypicality of nominal referents as a major parameter of their availability for human cognition, particularly in the mind of a speaker.

It seems reasonable to assume that human knowledge about the world is organized or structured in such a way that more frequent, animate, concrete, and/or prototypical concepts are easier to retrieve or make use of than less frequent, animate, concrete, and/or prototypical ones (Bartlett, 1932; Luchman, Luchman, & Butterfield, 1979; Neisser, 1967). It is also the case that the former are generally of more importance and relevance to everyday life and activities than are the latter (Brown, 1958, 1976; Brown & Lenneberg, 1954). As a result, more frequent, concrete, animate, and/or prototypical nominal referents seem to be better coded as lexical concepts in the conceptual knowledge (or the mental lexicon) than less frequent, concrete, animate, and/or prototypical nominal ones, so that the former are easier to learn, recall, recognize, or visualize than are the latter (Bock, 1982; Rosch, 1978; Rosch et al., 1976).

The present work is primarily concerned with the cognitive availability of nominal referents as the non-discourse determinants of the cognitive points of departure in the minds of speakers and thus of grammatical subjects. This leads to the following basic hypothesis of the present work: a speaker tends to choose the nominal referents of high cognitive availability as his cognitive points of departure and thus as grammatical subjects. It is also possible to define at least operationally the cognitive availablity of nominal referents in terms of their four inherent non-discourse properties: frequency, animacy, concreteness, and prototypicality.

On the basis of what has been discussed so far, it is possible to postulate the following five specific hypotheses concerning the psycholinguistic motivations for the choice of grammatical subjects.

Frequency Hypothesis. The more frequent a nominal referent is, the more likely it is to be chosen by a speaker as his cognitive point of departure and thus as a grammatical subject.

Concreteness Hypothesis. The more concrete a nominal referent is, the more likely it is to be chosen by a speaker as his cognitive point of departure and thus as a grammatical subject.

Animacy Hypothesis. The higher a nominal referent is in an animacy hierarchy, the more likely it is to be chosen by a speaker as his cognitive point of departure and thus as a grammatical subject.

Interaction Hypothesis. There may be interactive relations among some of these inherent properties of nominal referents, frequency, concreteness, and animacy, with regard to the selection of a grammatical subject.

Prototypicality Hypothesis. The more prototypical a nominal referent is as a member of a category, the more likely it is to be chosen by a speaker as his cognitive point of departure and thus as a grammatical subject.

According to the Frequency and Animacy Hypotheses, (1b) would be preferable over (1a), since the referent of *the woman* is more frequent and animate than that of *the shining wampum*. The Concreteness Hypothesis predicts that (2b) would be preferred to (2a), since the referent of *the locomotive* is more concrete that that of *the greatest human invention*.

This thesis tests these hypotheses in a psycholinguistic account for sentence subject selection in a series of experiments on sentence production. Three different experiments were designed and carried out, the first of which tested the Frequency, Concreteness, Animacy, and Interaction Hypotheses, and the second and third of which tested the Prototypicality Hypothesis.

It was found that there were significant interactions among the three inherent non-discourse properties, frequency, concreteness, and animacy, concerning the selection of grammatical subjects. First, the Frequency Hypothesis turned out to be true only with respect to human referents. Second, it was shown that the Concreteness and Animacy Hypotheses were confirmed only with regard to high frequency referents; that is, both hypotheses were rejected for low frequency referents. Finally, the prototypicality of nominal referents was shown to be a major non-discourse determinant of subject selection in a sentence.

In Chapter Four, these experimental results are compared with the results of several related psycholinguistic studies on the selection of grammatical subjects. Moreover, psycholinguistic implications of these results are discussed in terms of the possible psycholinguistic content for the notion of grammatical subject and for a nearly universal relative order of subjects before other nominal constituents.

2. REVIEWS OF RELATED STUDIES

It is necessary to review related studies in both the linguistic and psycholinguistic literatures in order to make clear the basic problem and the hypotheses of the present thesis. The following questions are of central importance for the review:

1. How plausible are the frequency, concreteness, animacy, and prototypicality of nominal referents as the parameters of cognitive availability (i.e., as non-discourse determinants)?
2. How have linguists and psychologists characterized the traditional notion of the speaker's cognitive point of departure as a primary psychological function of a grammatical subject?
3. What are linguistic implications of the present psycholinguistic work on subject selection in a sentence?

In subsequent sections, each of these three issues will be addressed in more detail by making reference to a number of related studies in the linguistic and psycholinguistic literatures.

2.1 *The Cognitive Availability of Nominal Referents*

A common sociocultural observation is that people have varying amounts and kinds of verbal and nonverbal experience with entities in both the external and the internal world as part of their so-called "episodic memory" (Lachman et al., 1979; Tulving, 1972). In other words, some entities are socioculturally more available and valid than are others. As a result, people may have more cognitive access to some nominal entities than to others in the sense that they can learn, recognize, or recall the former more easily (e.g., Brown, 1958, 1976; Paivio, 1969, 1971; Paivio & Begg, 1981).

Furthermore, some authors (e.g., Bock, 1982; Forster, 1976; Rosch, 1978) suggest that some nominal referents are better coded as lexical concepts in the mental lexicon than are others, with the former being easier to retrieve, recall, learn, or recognize than the latter. This line of reasoning seems possible, but such codability in the mental lexicon still remains open to some questions in the cognitive psychology literature (Anderson, 1980; Glass, Holyoak, & Santa, 1979; Lachman et al., 1979). For example, codability in the mental lexicon is hard to define precisely, primarily since the structure or organization of the mental lexicon has not yet been well explored, although a number of psychologists have proposed several models, for example Anderson and Bower's (1973) "human associative memory," Norman and Rumelhart's (1975) "active structural network," Schank's (1972) "conceptual dependency network," and so on.

Following these considerations, the present work starts with the assumption that nominal referents differ in their cognitive availability, but not in their codability in the mental lexicon, only in the sense that some are easier to process than others. It is much beyond the scope of this work to address the question of why some nominal referents are easier to process at a cognitive level than others. As the above-mentioned psychologists suggest, this is probably because some nominal referents are better coded as lexical concepts in the conceptual knowledge of humans than others.

The main focus of the present thesis is on the correlation between the cognitive availability of nominal referents and their likelihood of being the cognitive point of departure in the mind of a speaker and thus selected as a grammatical subject. To be more specific, it is possible that the speaker is likely to choose nominal referents of high cognitive availability as his cognitive points of departure rather than those of low cognitive availability, with the former being realized as grammatical subjects.

The specific question to be addressed in this section is, therefore, what inherent non-discourse properties of nominal referents contribute to their availability for human cognition. In other words, what inherent properties of the nominal referents are most representative of their availability for human cognition, especially human verbalization? As such, the present work is focused on the following four inherent non-discourse properties: frequency, concreteness, animacy, and prototypicality.

There is no doubt that people have more cognitive access to more frequent nominal referents than less frequent nominal ones, simply because the former can be experienced and observed more frequently in everyday life than the latter. Therefore, it is possible to argue that the frequency of nominal referents is one of the major parameters of their cognitive availability.

It has been shown in the psychology literature that people have less difficulty in processing more frequent verbal (e.g., nouns) and nonverbal (e.g., pictures) stimuli at a conceptual level than less frequent verbal and nonverbal ones. For example, more frequent words have been found to be much easier to learn and recall than less frequent ones (e.g., Earlhard, 1982; Forster & Chambers, 1973; Gregg, 1976; Gregg, Montgomery, & Castano, 1980; Hall, 1979). According to Noble (1952), words with high frequency can be viewed as being more meaningful than those with low frequency in the sense that the former can elicit more word associates than the latter (Underwood & Schulz, 1960). Finally, Oldfield (1963, 1965) and Oldfield and Wingfield

(1965) suggest that lexical entities should be stored in such a way that more frequent words can be made accessible more readily and can be retrieved much faster than less frequent ones (cf. Carrol & White, 1973; Lachman et al., 1979).

Forster (1976) stresses the frequency effects on word production, word meaning, word classification, and so forth as follows:

> By means of the so-called *lexical decision* experiment, we can establish the time required for lexical access to occur. In such an experiment, a letter sequence is presented for as long as the subject requires to classify the item as a word (in his vocabulary) or as a nonword as rapidly as possible. ...
>
> In such experiments, it is typically found that familiar words are classified in around 500 msec., but that nonwords require about 650 msec. Thus, rather than taking less time to classify, nonwords take substantially *more* time to classify than words. ...
>
> There is another critical fact that these theories are quite incapable of explaining in a natural way. The fact is that words which have a ralatively high frequency of occurrence are classified faster than words with a low occurrence frequency, e.g., *mildew, perspire, radiate*, although the latter are still perfectly familiar to the subjects of the experiments. ... (original emphases, pp. 260-263)

From this, it is clear that the frequency of words is a major parameter governing their ease of cognitive processing.

As one possible explanation of frequency effects, Forster suggests that lexical entities should, in principle, be ordered or organized in terms of their relative frequencies of usage, so that people may have more access to frequent words than infrequent ones. This is also consistent with Oldfield's (1965) view of the mental lexicon.

The present thesis is not concerned with why such frequency effects on lexical decisions are possible, but instead with a well-established experimental fact that frequent nouns or their referents must be of higher availability for human cognition. Therefore, it is possible to assume that frequent nominal referents are more available in the mind of a speaker than infrequent ones.

It is on the basis of this line of reasoning that the Frequency Hypothesis is proposed in this thesis. It states that the more frequent a nominal referent is, the more likely it is to be chosen by a speaker as his cognitive point of departure and thus as a grammatical subject.

In this work, the frequency of nominal referents is operationally defined in terms of the individual frequencies of occurrence of the corresponding nouns. It is important to mention that the present thesis is not concerned with the class or category frequencies of nominal referents, such as humans, animals, and inanimate entities. This is because the category frequencies of nominal referents are correlated with the animacy hierarchy; in other words, the category frequency and animacy of nominal referents are not independent. On the other hand, individual frequencies of nominal referents are independent of their animacy. It is, for example, easy to find human nouns whose frequencies of usage are as low as those of some animals or inanimate entities. For example, according to Thorndike and Lorge (1944), the individual frequency counts of human nouns, *idler* and *mentor*, are two and one per million words; those of non-human nouns, *weevil* and *wampum*, are two and one per million words.

It can be argued that frequency is not an inherent property of nominal referents; rather it generally has to do with the usage or occurrence of the nominal referents. However, it is clear on the basis of the above discussion that frequency must be viewed as a non-discourse contributor to the cognitive availability of nominal referents. In this sense, the present thesis treats the frequency of usage of nominal referents as if it were an inherent non-discourse property.

The concreteness of nominal referents is proposed here as the second parameter of availability for human cognition. People can have both verbal and nonverbal direct experience with concrete nominal referents or objects, whereas they do not have direct nonverbal experience with abstract nominal referents. In other words, the human experience with abstract entities is much more limited than with concrete ones. From this, it follows that people may have more cognitive access to concrete nominals in that they can process them in both a verbal and nonverbal manner, while they can process abstract nominals in a verbal manner only (e.g., Kieras, 1978; Paivio, 1971; Paivio & Begg, 1981).

It has been shown in the psychology literature that concrete verbal stimuli (e.g., concrete nouns) are much easier to process at a cognitive level than abstract verbal stimuli (e.g., abstract nouns). In other words, people can perform with concrete verbal stimuli in a variety of cognitive tasks much better than with abstract verbal stimuli. For example, people can learn, recall, and recognize concrete noun pairs (e.g., *tree-stone*) much more easily than abstract noun pairs (e.g., *honesty-fact*) (Paivio, 1971, Paivio & Begg, 1981; Paivio, Yuille, & Rogers, 1967; Reynolds & Paivio, 1968; Yuille & Paivio, 1967). These experimental results are partially due to the fact that concrete and abstract nouns are different in their inherent capacity to evoke conceptual images. In other words, concrete nouns can arouse conceptual images much more easily than abstract nouns; and such images may help subjects to learn and recall concrete nouns as conceptual mediators.

Kieras (1978), for example, argues that image formations of concrete and abstract concepts are based on different information processes. He states:

> In the case of a concrete word, the process of forming an image consists of finding the concept node in memory and retrieving the attached perceptual description. The subjects can then signal the awareness of an image and rate its ease of formation or vividness.

> In contrast, for abstract words, the image formation process is rather different, since there is no perceptual description attached to an abstract concept node. For example, since *justice* is abstract and is not an object there is no perceptual description of *justice* to simply retrieve, as is the case for concrete words. However, images *related* to abstract words can be devised. For some abstract words, there are conventional images, such as that of a blindfolded woman with scales for *justice* ... If there are no conventional images, further search into the memory network may lead to related concrete concepts. For example, further search will find the proposition that justice is often dispensed by a *judge*; this concept is concrete and has a perceptual description. Together with other related information, a detailed image for *justice* can be constructed, such as a stern-faced John Sirica sentencing a group of Watergate defendants. (original emphases, pp. 538-539)

It is evident that image formation of abstract concepts needs additional memory processes, thereby taking more time than that of concrete ones. In this sense, concrete concepts may be processed more easily at a cognitive level than abstract ones, with the former being of higher cognitive availability than the latter.

Furthermore, concrete sentences (e.g., *Tom hit John*) were found to be much easier to learn, recall, and recognize than abstract sentences (e.g., *A good deed is a human virtue*) (Moeser, 1974; Paivio, 1971; Paivio & Begg, 1981; Walter & Fox, 1981). It has also been argued that concrete nominal referents or concepts are much better coded, organized, and categorized within the conceptual knowledge (or semantic memory) of humans in terms of their categorical hierarchical interconcept relations (e.g., "paradigmatic" and "syntagmatic" relations) (Hampton, 1981; Miller &

Johnson-Laird, 1976; Porzig, 1938; Quillian, 1968; Richards, 1976; Rosch, 1973, 1974, 1975; Rosch et al., 1976).

Taken together, these studies suggest that concrete nominal referents should be viewed as being more available for human cognition than abstract nominal referents in the sense that the former can be processed much more easily at the conceptual level than the latter. Particularly, Paivio and Begg (1981) argue for the superiority of concrete words over abstract ones in human cognition as follows:

> A strong generalization is justified by the available research evidence: the image-evoking value of words is the best predictor of performance in most verbal learning and memory task. ... High-imagery words are easier to remember than low-imagery words even when variables are controlled, and the effects of other semantic attributes are generally weak or nonexistent when imagery is controlled. In one study (Paivio, 1968), for example, over twenty characteristics of words were correlated with paired-associates learning and free-recall scores of the same items. Only a few of the attributes significantly predicted the memory scores: the highest correlation was between rated imagery values and recall scores. The other predictors of memory performance, in decreasing order of predictive power, were ratings of the vividness of imagery, a reaction-time measure of the speed with which the word arouses images, and rating of concreteness and tangibility. In fact, a factor-analytic study showed that all of these variables correlated highly with each other, constituting what was called in the study an imagery-concreteness factor. (pp. 185-186)

It is reasonable, therefore, to assume that high-imagery (i.e., highly concrete) nominal referents are a reliable predictor of cognitive availability.

This thesis is not concerned with why there is a positive correlation between the imagery-concreteness of nominal referents and their ease of processing, but instead with a well-established experimental fact that concrete nominal referents are of higher availability for human cognition than abstract ones. Thus, it is reasonable to assume that concrete nominal referents are more available in the minds of speakers than abstract ones.

It is on these grounds that the Concreteness Hypothesis is proposed here. It predicts that the more concrete a nominal referent is, the more likely it is to be chosen by a speaker as his cognitive point of departure and thus as a grammatical subject.

The animacy of nominal referents/concepts is offered as the third parameter of their availability for human cognition. Traditionally, it has been maintained that people perceive and think about their verbal and nonverbal experience with their external world primarily from their egocentric point of view (e.g., Bock, 1982; Hörmann, 1981; Piaget, 1926; Lyons, 1977). This implies that they have an attentional bias towards humans and human-related entities in thinking, perception, and so forth. Consequently, the higher nominal referents are in an animacy hierarchy, the easier they are to process at a cognitive level.

There seem to be two basic lines of argumentation for such egocentric human cognitiion. The first has to do with the well-known notions of "egocentrism" and "animism," which Piaget first introduced into developmental psychology (e.g., Flavell, 1963; Piaget, 1926, 1929, 1970). The second is the so-called "anthropocentric" nature of human thinking, which simply means that people are more concerned with people themselves than with non-humans.

Piaget consistently maintained that two major characteristics of young children's cognition were their egocentrism and animism. By "egocentric" he generally meant that the young child perceives his external world primarily from his own point of view, failing to take into consideration the points of view of others. On the other hand, the animistic aspect of young children's cognition is that they tend to regard inanimate objects as living and conscious. As a result, they fail to recognize certain semantic anomalies, such as in the following sentences: *Tables are living, Stones died, Bicycles are crying*, and so on (Russell, 1940a, 1940b, 1942; Russell & Dennis, 1939).

These two characteristics may also be found, to some extent, in adult human cognition, although they generally tend to decrease with age (Flavell, 1963). Rohrman (1970), for example, found that subjects could recall animate nouns more easily than inanimate ones (cf. Holland & Rohrman, 1979). It seems reasonable, therefore, to assume that the egocentric characteristic of human cognition is not restricted to the thinking and perception of young children.

There is another line of argument for egocentric human cognition, which has to do with the anthropocentric nature of human thought (Givón, 1976; Krupa, 1982; Lyons, 1977; Zubin, 1979). This position maintains that people tend to be more concerned with human interests and values than others. For example, Lyons (1977) points out linguistic aspects of anthropocentrism in human thinking as follows:

> Independently of these context-dependent considerations, however, we may be assumed, as human beings, to be more interested in persons than we are in animals, to be more interested in animals than we are in inanimate entities, and so on. It follows that in any one clause utterance in which reference is made both to a person and to an animal or inanimate entity, the expression referring to the person will be made thematic, unless there are special reasons for doing otherwise. (pp. 510-511)

(Lyons simply follows Halliday (1967, 1970, 1979) with regard to the notion of theme, which will be discussed in detail in the next section).

Following these considerations, the present work assumes that animate nominal referents, especially humans, are of higher availability for human cognition than inanimate nominal ones. Furthermore, it may be assumed that animate nominal referents may be more available in the minds of speakers than inanimates.

It is on the basis of this line of reasoning that the Animacy Hypothesis is proposed in this work. It states that the higher a nominal referent is in the animacy hierarchy, the more likely it is to be chosen by a speaker as his cognitive point of departure and thus as a grammatical subject.

According to Comrie (1981), however, it is difficult to define an explicit animacy hierarchy from a linguistic point of view. This is because some language-specific hierarchies of animateness of nominal referents do not always coincide with a conceptual animacy hierarchy, namely: human > animate non-human > inanimate.

For example, Comrie points out that in Navaho, nominal referents such as wind, rain, lightning, and so on are classified as being closer to animate nominal referents than to inanimate nominal referents. Citing Navaho translations of *The lightning killed the horse* and *Old age killed the horse*, he goes on to state that two noun phrases, *the lightning* and *the horse*, are considered to be so close in the animacy hierarchy of Navaho that two different verb prefixes, *yi-* and *bi-*, can be used as free variants. However, in the second Navaho sentence, only one verb prefix, *bi-*, can be used. Comrie attributes this restriction to the markedness of the word order in the second sentence whose first subject noun phrase, *Old age*, is much lower in the animacy hierarchy than the second noun phrase, *the horse*. If this is true, the two noun phrases, *The lightning* and *The old age*, are ranked differently in the animacy hierarchy in Navaho, although the referents of both are conceptually inanimate.

It is interesting, however, to note that such nominal referents as wind, rain, lightening, and so on can also be viewed as being more agentive than such inanimate nominal referents as old age, truth, and so on. This is because the former can, although inanimate in the conceptual animacy hierarchy, be potential sources for some actions, whereas the latter can not (Chafe, 1970; Cruse, 1973; Morley, 1983).

Tomlin (1979) attempts to reconcile the two seemingly independent linguistic and conceptual animacy hierarchies. He defines the notion of animateness of noun phrases on the basis of both their conceptual animacy hierarchy (i.e., human > animal > inanimate) and semantic case hierarchy (i.e., agent > beneficiary, ..., dative > patient). On the grounds of the twofold animacy hierarchy, Tomlin goes on to predict: human agent > ... > inanimate agent > ... animate patient > .... > inanimate patient, where "A > B" indicates that A is higher in the hierarchy than B. It is important to note that an inanimate agent entity can be viewed as more animate than an inanimate patient, although both are conceptually inanimate entities, since the former is higher in the semantic case hierarchy than the latter.

Returning to the Navaho example, it is obvious that, according to Tomlin's animacy hierarchy, inanimate nominal referents, such as wind, rain, and so on, can be judged to be more animate than other inanimate nominal referents such as old age, since the former fall in the inanimate agent category, while the latter are inanimate sources or instruments.

The present paper is not, however, concerned with such formal treatment or definition of the notion of animacy of nominal referents, but instead with the conceptual animacy hierarchy: human > animal > inanimate (Dewart, 1979; Fillmore, 1977; Harris, 1978; Kuno & Kaburaki, 1977). It is on the basis of the conceptual animacy hierarchy that the Animacy Hypothesis of the present thesis is proposed and discussed.

At this point, one point must be made about the relationship between the concreteness and animacy of nominal referents. These two can be considered to be either dependent on or independent of one another. The dependent aspect is that concrete referents, human and animate non-human, are higher in the animacy hierarchy than abstract ones. The independent aspect is that human, animate non-human, and object referents are ranked differently in the animacy hierarchy, although all of them are classified as concrete. Following these considerations, the Concreteness and Animacy Hypotheses will be tested both dependently and independently as explained in Section 3.1.

Finally, this work is concerned with the prototypicality of nominal referents as the fourth major parameter of availability for human cognition. It often happens that people have more difficulty naming some nominal referents than others. It is also possible that people may disagree about what to call one particular nominal referent. What these commonplace verbal experiences suggest is that there is some *vagueness* in the denotative relationships between some nominal referents or concepts, i.e. "the signified," and the corresponding name-words, i.e. "signifies," (Black, 1949; Fillmore, 1975; de Saussure, 1916; Lakoff, 1972; Lyons, 1977; Schmidt, 1974).

Black, for example, describes denotative vagueness as follows:

> A typical example of vagueness is described. A symbol's vagueness is held to consist in the existence of objects concerning which it is intrinsically impossible to say either that the symbol in question does, or does not, apply. The set of all objects about which a decision as to the symbol's application is intrinsically impossible is defined as the "fringe" of the symbol's field of application. It is claimed that all symbols whose application involves the recognition of sensible qualities are vague, ... (pp. 28-29)

It is clear that people may have great difficulty in naming referents in a certain fringe area and that they may not agree on what to call them. It is in this sense (i.e., naming difficulty and interpersonal disagreement on naming) that some nominal referents can be judged to be more prototypical of a category than other nominal referents. That is, the former may be in the fringe area, while the latter may not (Brown, 1958, 1976; Brown & Lenneberg, 1954; Lachman, 1972; Lachman et al., 1979; Rosch, 1973, 1974, 1975, 1978; Smith & Medin, 1981).

A typical example of this kind of phenomenon is color terms: it has been shown that English-speaking people have a great deal of difficulty in naming color stimuli whose wavelengths fall in the boundary regions (on the fringe) between those of two adjacent colors such as "blue" and "green," "green" and "yellow," "yellow" and "orange," and "orange" and "red" (Berlin & Kay, 1969). This is primarily because they can not easily determine to which of the two conceptual categories each of the color stimuli belongs. They may hesitate, disagree among themselves, and use different phrases or combinations of names, for example, "rusty orange," "desert red," "dried orange peel," "reddish orange," "orange-like red," and so on. The point is that some color stimuli are easier to classify and name than other colors, and can therefore be viewed as being closer to a prototype. It is also important to note that non-prototypical stimuli may be named in terms of phrases rather than single noun words; they are not codable in a linguistic sense (e.g., Brown, 1958, 1976; Chafe, 1977).

Labov (1973) made an interesting experimental attempt to measure directly the denotative vagueness of English words "cup" and "bowl." He provided the subjects with a series of pictures of cup-like objects, where the items numbered 1-5 were designed to implement the increment of the ratios of the width to the depth; the ratios of the five "cups" were 1.0, 1.2, 1.5, 1.9, and 2.5; and the depth was constant across all the items. The subjects were simply asked to name each of these cup-like objects.

The analysis was focused on whether the subjects used *cup* or *bowl* as the head nouns in their naming. It was found that the percentages of "cup" subjects were 100%, 100%, 85%, 35%, and 20% for items 1 through 5, respectively, while those of "bowl" subjects 0%, 0%, 0%, 0%, and 25%. This implies that items 4 and 5 are much more difficult to categorize as either cups or bowls and therefore much more difficult to name than items 1-3.

It follows that the prototypicality of items 1 through 3 is much higher than that of items 4 and 5. More importantly, this can be attributed to the experimental manipulations whereby items 1 through 3 are much easier to categorize as the instances of cups than are items 4 and 5, since the shapes of the former are more prototypical of cups than those of the latter. That is, typical instances of cups are easier to process than atypical ones (Chafe, 1977; Rosch, 1973, 1974, 1975; Rosch et al., 1976).

Another well-known example of prototypicality relates to the fact that one can generally make faster "truth" responses to such target sentences as *A robin/sparrow is a bird*, than to such target sentences as *A chicken/duck is a bird*. (Note that the truth values of both sentences are the same.) This is because robins and sparrows are among the most prototypical birds, thereby being easy to retrieve and process at a cognitive level, while chickens and ducks are not (e.g., Caramazza & Brones, 1980; Rosch, 1973, 1978; Smith & Medin, 1981).

This thesis is not concerned with why the prototypicality of nominal referents plays such an important role in human cognition (cf., Berlin & Kay, 1969; Rosch, 1972), but with the experimental finding that prototypical nominal referents are apparently of high availability for human cognition. It is in turn possible to assume that relatively prototypical nominal referents are more available in the minds of speakers than are non-prototypical ones.

It is on these grounds that the Prototypicality Hypothesis is proposed in this work. It predicts that the more prototypical a nominal referent is, the more likely it is to be selected by a speaker as his cognitive point of departure and thus as a grammatical subject.

It is true that prototypicality is not inherent in nominal referent in the same sense that concreteness and animacy are. That is, prototypicality is by nature associated with nominal referents as instances of a certain category; on the other hand, concreteness and animacy are both associated with nominal referents as individuals. This difference should not be crucial, as long as prototypicality can be viewed as an independent non-discourse contributor to the cognitive availability of nominal referents, which has been verified above. Therefore, it seems justifiable to treat the prototypicality of nominal referents as an independent inherent non-discourse property.

What has been discussed so far can be summarized in the following three points. First, it has been shown in the psychology literature that more frequent, concrete, animate, and/or prototypical nominal referents are easier to process (i.e., learn, recall, recognize, etc.) at a cognitive level than less frequent, concrete, animate, and/or prototypical nominal ones. As stated earlier, this is *presumably* because the former are better coded as lexical concepts in the mental lexicon than are the latter. It is much beyond the scope of this thesis, however, to seek to relate the cognitive availability of nominal referents/concepts with certain specific structures or organizations of the mental lexicon. Second, it seems reasonable to hypothesize that the frequency, concreteness, animacy, and prototypicality of nominal referents may be the major parameters for defining availability in human cognition, particularly in the mind of a speaker. Third, it seems possible to hypothesize that more frequent, concrete, animate, and/or prototypical nominal referents are more available in the mind of a speaker than less frequent, concrete, animate, and/or prototypical ones, with the former being more likely to be chosen by the speaker as his cognitive point of departure and thus as a grammatical subject. The primary purpose of the present work is to investigate the cognitive availability of nominal referents as the non-discourse determinants of the cognitive point of departure in the minds of speakers and thus of grammatical subjects.

2.2 Grammatical Subject as Cognitive Point of Departure

A number of linguists and psycholinguists agree that a primary psychological function of a grammatical subject is to refer to the cognitive point of departure in the mind of a speaker. However, there is less agreement about how to characterize and investigate the traditional notion of the speaker's cognitive point of departure as such.

Actually, the following purely descriptive notions of the speaker's cognitive point of departure can be enumerated as part of the linguistic explanation of subject selection in a sentence: "theme (as 'given' information or 'degree of communicative dynamism')" (e.g., Firbas, 1964; Mathesius, 1964), "theme (as the sign of a speaker's concern)" (e.g., Allerton, 1978; Halliday, 1967), "a speaker's perspective (or point of view)" (e.g., Allerton, 1979; Dik, 1978), "a speaker's focus of empathy" (e.g., Kuno & Kaburaki, 1977; Itagaki, 1982; Ramson, 1977), and "topic (as that which a sentence is about)" (e.g., Givón, 1983; Hockett, 1958; Li & Thompson, 1976).

Alternatively, a number of psycholinguists have proposed the following notions of a speaker's cognitive point of departure as part of a psycholinguistic account of subject selection in a sentence: "perceptual salience (e.g., size, shape, etc.)" (e.g., Clark & Chase, 1974, Flores d'Arcais, 1973) "figure (as opposed to 'ground')" (e.g., Osgood & Bock, 1977; Talmy, 1978), "degree of importance" (e.g., Johnson-Laird, 1977), "a speaker's reference point of sentential construction" (e.g., Ertel, 1977), and "a speaker's focus of interest" (e.g., Zubin, 1979).

These various proposals imply that some particular nominal referents are more likely than others to be chosen by speakers as their cognitive points of departure and thus as grammatical subjects. The question is then how to characterize the preferences for some particular referents over others as the speaker's cognitive points of departure. It is proposed here that one way to characterize them is in terms of four inherent non-discourse properties of nominal referents: frequency, concreteness, animacy, and prototypicality.

In the following two sections, this point will be further justified in conjunction with a detailed review of each of the linguistic and psycholinguistic notions of the speaker's cognitive point of departure. For the sake of convenience, the linguistic and the psycholinguistic notions of cognitive point of departure will be reviewed separately.

2.2.1 Linguistic Studies of Cognitive Point of Departure

It is the Prague linguists, such as Danes, Firbas, and Mathesius who first explored the traditional notion of the speaker's cognitive point of departure within the framework of descriptive linguistics, more specifically within the framework of the so-called "functional sentence perspective" or "contextual sentence organization" (Vachek, 1966).

Their main efforts were focused on (a) how to characterize the notion of cognitive point of departure as a discourse relation within a sentence; and (b) how to relate it to linguistic phenomena, especially grammatical subjects and word order phenomena. They have indeed elaborated the notion of the speaker's cognitive point of departure as a purely descriptive discourse relation, which they called "theme" within a sentence.[2]

Mathesius (1964), originally published in *Actes du Premier Congres International de Linguistes a La Haye*, 1928, pp. 56-63, first proposed to refer to the traditional notion of cognitive point of departure by the linguistic term "theme," suggesting that a primary function of grammatical subjects in Modern English is to refer to the theme of the sentence. He stated:

> In languages with developed verbal systems there very often appears a vacillation between two

---

[2] The Prague linguists do not use the term "cognitive point of departure," but instead the term "psychological subject."

> different conceptions of the grammatical subjects, that of the doer of the action expressed by the predicative verb and that of theme of enunciation contained in the predicate.[1] (p. 61)

In the note, he stated: "[1]By the terms 'theme' and 'enunciation' I mean what is usually called the psychological subject and psychological predicate" (p. 67).

According to Vachek (1966) and Firbas (1964, 1966), Mathesius goes on to define the notion of theme as "given" information somehow recoverable from the preceding discourse or situational setting. This suggests that he is primarily concerned with the discourse aspects of a speaker's cognitive point of departure and then of a grammatical subject. As discussed in Chapter One, it seems true that given information is more available in the mind of a speaker than new information, with the former being likely to be chosen by the speaker as his cognitive point of departure and thus as the grammatical subject (cf. Bock, 1982; Itagaki, 1982).

Firbas (1964, 1966, 1974) agrees with Mathesius that the primary psychological function of a grammatical subject is to refer to the cognitive point of departure of the speaker (i.e., theme). He does not agree, however, with Mathesius as to how to define the notion of theme. Firbas first claims that sentence elements differ in their communicative contributions to the on-going discourse. In order to capture this, he then proposes the notion of "communicative dynamism" (hereafter CD), which he defines as follows: "By degree of CD carried by a linguistic element, I mean the extent to which the element contributes towards the development of the communication" (Firbas, 1974, p. 19). He goes on to define the notion of theme (i.e., psychological subject) as sentence elements with the lowest degree of CD.

Firbas also proposes that the CD of each sentence element is a function of three basic variables: positions in a sentence, contextual dependency, and semantic content. First, as the most natural distribution of CD, the degree of CD increases from sentence-initial through sentence-middle to sentence-final elements. Second, given information has for the most part the lowest degree of CD, with new information having the highest degree of CD. Third, sentence elements are by nature different with respect to their semantic importance for communication. Firbas, for example, assumes the following hierarchy of the three semantic elements involved in action events: Goal > Action > Agent, where "A>B" means that A is communicatively more important than B. The final CD distribution of each sentence is simply the result of the interplay of these basic three variables. It is important to note that (a) the first is related to grammatical word order, for example subject - verb - object in English; (b) the second has to do with the discourse aspects of sentence elements (Chafe, 1976); and (c) the third seems to be related to the non-discourse aspects of sentence elements.

What is crucial here is that Firbas attempts to capture both discourse and non-discourse aspects of nominal referents as the determinants of the theme of a sentence as the cognitive point of departure of the speaker, which is different from Mathesius' approach to the notion of theme. In addition, the non-discourse determinant, the third variable, seems to be related to the notion of cognitive availability of nominal referents, particularly to the animacy of the referents. As another example of the third variable, Firbas provides the following semantic hierarchy: Possessed > Possession > Possessor. The Possessor and Agent are typically animate semantic elements, while the Goal and Possessed are typically inanimate semantic elements, with the former being likely to have the lowest degree of CD and then to be the theme of a sentence. It is possible to view the discourse determinant (i.e., the second variable of CD) of the cognitive point of departure as primary (Mathesius, 1964), but as suggested by Firbas (1964, 1966, 1974), some inherent properties of nominal referents, especially animacy, should not be ignored as the non-discourse determinant of the cognitive point of departure and thus of the grammatical subject.

It is clear that the Prague linguists view the traditional notion of the speaker's cognitive point of departure as the primary psychological function of a grammatical subject, but they do not agree as to how to characterize or define it, particularly in terms of the discourse and non-discourse aspects. In particular, Firbas suggests that the cognitive point of departure in the mind of a speaker should be determined not only by the discourse aspects, but also by the non-discourse aspects of nominal referents (cf. Mathesius, 1964).

Another well-known elaboration of the notion of the speaker's cognitive point of departure has been proposed and discussed by Allerton (1978) and Halliday (1967, 1970, 1979). They first claim that the speaker's cognitive point of departure may exist at two linguistic levels: tone groups and clauses, where the tone groups for the most part coincide with the clauses. From the viewpoint of sentence production, this may be taken to mean that a speaker may have more than one independent cognitive point of departure in the planning of the sentence, depending on the number of tone groups and clauses in the sentence.

From a linguistic point of view, Halliday refers to the speaker's cognitive points of departure for tone groups and clauses as "given" information and "theme," respectively. Halliday's (1979, p. 67) example is helpful for the understanding of these points:

```
    // the MOON // was shining SKULKILY //
          Theme / -------- Rheme -------
           New  / -- Given -- / - New -
```

where "//" indicate the boundaries between tone groups, which Halliday views as a "basic information unit" in discourse (cf. Fronek, 1983) and the capitalized phrases indicate the tonic positions in the tone groups.

A few additional points must be made here. First, at the levels of the tone groups and the clause, the action, *was shining*, and the referent of *the MOON*, respectively, function as the cognitive points of departure of the speaker. Second, according to Halliday, the "given" information is not an obligatory element in the tone group, as in the case of the first tone group of the above example; however, from a communicative point of view, the "new" information must be in the tone group. Third, the grammatical subject of the clause serves the theme function, i.e., the cognitive point of departure of the speaker. Furthermore, suppose that the above example is uttered as a single tone group, for example // The moon was shining SKULKILY //. In this case, the grammatical subject, *The moon*, functions as both the theme and part of the given information, i.e. two independent cognitive points of departure of the speaker.

Halliday (1979) describes the main motivation for the two independent aspects (i.e., theme and given information) of the traditional notion of cognitive point of departure as follows:

> So there is a peak of prominence at the beginning, which is the Theme: and another peak of prominence, usually at the end, which is the focus of information or, simply, the New. The two are different in meaning. The theme is speaker-oriented; it is *the speaker's signal of*

> *concern, what it is that he is on about* - he may even make this explicit, by starting 'as far as ... is concerned'. The new is hearer-oriented (though still, of course, SELECTED [an original emphasis] by speaker); it is the speaker's presentation of information as in part already recoverable to the hearer (the Given) and in part not recoverable (the New). The two types of prominence are independent of each other. But both contribute to the 'texture' to fashioning the fabric of the text. (emphasis added, p. 68)

If this is so, the two different views of a speaker's cognitive point of departure (i.e., theme and given information) do not seem to be reconcilable, since they reflect two independent psychological decisions which speakers have to make during their planning of sentences. One has to do with "focus of concern"; the other with assumptions about hearers' knowledge (i.e., given vs. new information). For example, in the following clause and tone group // Tom is going to TOKYO //, the speaker is assumed to have made the following two psychological decisions about his cognitive points of departure: he decided to talk about the referent of *Tom*; and he assumed that his hearers already knew that Tom was going somewhere, but they did not yet know where Tom was going. Halliday has characterized these two decisions as two linguistic discourse relations within a sentence, theme vs. rheme and given vs. new information, respectively.

It is not the principal concern of the present work to determine whether or not Halliday is right in arguing for the two independent aspects of theme and given information of the traditional notion of a speaker's cognitive point of departure. Instead, this work is concerned with the fact that he views the primary psychological function of a grammatical subject as referring to the cognitive point of departure, which he defines as "[the speaker's] signal of concern, what it is that he is on about" (Halliday, 1979, p. 68).

An important question must be addressed concerning Halliday's notion of theme: Why are some nominal referents more likely to be chosen by a speaker as his focus of concern than others? The major determinant of the focus is, of course, the speaker's considerations of the on-going discourse, as reflected, for example, in the distribution of given-new information. It is also possible that other inherent non-discourse properties of nominal referents are involved in the speaker's choice of the focus of concern (i.e., theme as his cognitive point of departure).

As an example, suppose that (a) a speaker chooses *The man was hit by a car*, whose theme is the referent of *The man*, rather than *A car hit the man*, whose theme is that of *A car*; and (b) the referent of *the man* is given information and that of *a car* is new information. One reasonable explanation is that the speaker chooses the given information, *the man*, as the theme (i.e., his focus of concern and then the cognitive point of departure) rather than the new information. There is, however, another possible explanation: the speaker chooses the referent of *the man* rather than that of *a car* as the theme of sentence, since the former is a human and the latter is an object. It is conceivable that speakers are for the most part more concerned with given information and humans than new information and non-human entities, respectively (Lyons, 1977). Furthermore, it is possible that both the discourse (e.g., given information) and non-discourse (e.g., animacy) aspects of the referent of *The man* contribute together to its likelihood of being the theme in the sentence.

It is clear that the choice of theme has to do with both the discourse and non-discourse aspects of nominal referents, if the theme of a sentence functions as the focus of concern of the speaker. The choice of theme seems to be determined not by the discourse factors alone, although Halliday proposes the notion of theme as a discourse relation. The present work is concerned with the inherent non-discourse properties of nominal referents (i.e., cognitive availability) as the major determinants of the cognitive point of departure of a speaker and then of the grammatical subject.

Another popular way of characterizing the traditional notion of the speaker's cognitive point of departure is in terms of the perspective (or point of view) from which the speaker presents the state of affairs. As one communicative action, the speaker chooses one particular participant (i.e., nominal referent) in the state of affairs as his cognitive point of departure, describing the state of affairs from the perspective of that chosen participant (Allerton, 1979; Dik, 1978, 1980, 1983; Fillmore, 1977; Itagaki & Prideaux, 1983). In addition, these authors claim that the primary function of a grammatical subject is in fact to specify the speaker's perspective, i.e., his cognitive point of departure.

This can be readily seen from the following statements by Dik:

> ... we defined the Subject as that constituent which refers to the entity which is taken as a point of departure for the presentation of the state of affairs in which it participates. (Dik, 1978, p. 87)

and

> It should be noted that, although the term 'syntactic function' is used (mainly for traditional reasons), FG [Functional Grammar] gives a semantic interpretation to the functions Subject and Object. The semantic value of these functions is not judged to contribute to the definition of the state of affairs designated by the predications as such, but to *the particular way in which this state of affairs is presented, i.e. the point of view from which it is described* (emphasis added, Dik, 1980, p. 13).

The point is that such a choice of "the particular way" may not only be viewed as a descriptive notion in linguistic theories, but also as a pscyhological aspect of the way speakers produce sentences.

The perspective or point of view of a speaker is referred to as "empathy focus" by Kuno (1976), Kuno and Kaburaki (1977), and Ramson (1977). They suggest that the speaker tends to place his empathy focus with one particular participant in the state of affairs, describing the state of affairs from the point of view of that participant. It seems that there is no essential difference between the notion of empathy focus and the notion of a speaker's perspective. Thus, the present thesis will deal with these two as the same linguistic device to characterize the traditional notion of the speaker's cognitive point of departure as the primary psychological function of a grammatical subject.

The psycholinguistic evidence for this view of a grammatical subject has been demonstrated by Itagaki and Prideaux (1983), who had subjects listen to short stories and then rewrite them from the perspective of one particular character in each of the stories. In each of the original stories, two characters were equally prominent, so that both of them could be taken to be the main character. In retelling the stories, the subjects tended to choose one particular character (i.e., the target of the perspective) significantly more often than the other as agent and subject. It must be noted that, in the original stories, the frequency with which the two characters occurred in the grammatical subject positions in the clauses were manipulated so as to be almost equal.

What is relevant to this work is that some nominal referents may be more likely to be chosen as the target of a speaker's perspective than others. Fillmore (1977) and Kuno and Kaburaki (1977), for example, suggest that the choice of the target of perspective should have to do with the salience, animacy, and definiteness of nominal referents. Specifically, the more salient, animate, and/or definite a nominal referent is, the more likely it is to be chosen by a speaker as the target of perspective, which is in turn realized as a grammatical subject. From this point of view, it is important to recognize that the choice of the target of perspective is made by a speaker on the basis of both a discourse feature (e.g., definiteness) and some non-discourse features (e.g., salience and animacy) of nominal referents.

Itagaki (1982) tested the hypothesis that the higher a nominal referent is in an animacy hierarchy, the more likely it is to be the target of the speaker's perspective and thus to be a grammatical subject in a sentence. The experimental results were inconclusive on the relationship between the animacy of nominal referents and their likelihood of being the targets of perspective and then grammatical subjects. This is partly because he did not take into account some inherent properties other than animacy, which seem to be involved in the choice of the target of perspective and then grammatical subject. Following these considerations, the present work is concerned with the frequency, concreteness, and prototypicality as well as animacy of nominal referents as important non-discourse determinants of the choice of the target of perspective (i.e., the cognitive point of departure) and accordingly of the grammatical subject.

Finally, a number of American linguists have characterized the traditional notion of the speaker's cognitive point of departure as a discourse relation of "topic," namely "that which the sentence is about." (Creider, 1979; Bolinger, 1981; Givón, 1976, 1977, 1979a, 1979b, 1983; Hockett, 1958; Li & Thompson, 1976). They generally agree that a grammatical subject serves primarily to refer to the topic of a sentence, i.e., the cognitive point of departure in the mind of a speaker. Hockett, for example, introduced the notion of topic as follows:

> The most general characterization of the predicate construction is suggested by the term "topic" and "comment" for their ICs [Immediate Constituents]: the speaker announces a topic and then says something about it. Thus *John / ran away; That new book by Thomas Gurnsey / I haven't read yet*. In English and the familiar languages of Europe, *topics are usually subjects; and comments are predicates*: so in *John / ran away*. (emphasis added, p. 201)

From this, it follows that the primary psychological function of the grammatical subject is to refer to the topic of the sentence, which serves as the cognitive point of departure in the mind of the speaker. Moreover, the topic of the sentence necessarily occurs prior to the comment.

On the basis of his typological studies, Creider (1979) argues that the association of grammatical subjects with discourse topics is nearly universal. He states:

> We may summarize the current state of

> knowledge of the association of linear ordering and discourse factors across languages in terms of three major strategies, each correlated with a major syntactic order type. Languages that treat initial position as topical and final position as focusing are SVO (English, Spanish, Czech, Russian). Languages that treat initial position as topical and preverbal position as focusing are verb-final (Hungarian, Quechua). Finally, languages that treat initial position as focusing and final position as topical are verb-initial (Nandi, Tagalog, Malagasy). These latter languages always have a means of reversing this order to produce sentences that have initial topics. This reversed order is found in discourse context where the topic is not known or predictable from the preceding context... (p. 19)

If this is so, the topical function of the grammatical subject should be regarded as nearly universal. It is, of course, true that languages are different with regard to the extent to which such topical function is grammaticalized as the grammatical subjects of basic sentences (Li & Thompson, 1976).

The next question is how the topicality of noun phrase arguments in a sentence can be determined. Givón (1976) suggests that the following four hierarchies should be postulated as such: 1) human > non-human, 2) definite > indefinite, 3) more involved participant > less involved participants (i.e., Agent > Dative > Accusative), and 4) 1st person > 2nd person > 3rd person, where "A>B" means that A is more likely to be chosen as the topic (i.e., the cognitive point of departure of a speaker) than B.

Clearly, these four are composed of both discourse and non-discourse aspects of nominal referents. That is, the second can be conceived of as a discourse determinant of topic (i.e., given-new information structure), while the first, third and the fourth are non-discourse determinants (i.e., inherent properties of nominal referents) of topic. Furthermore, the first, third, and fourth may be related to the animacy of nominal referents, although the fourth is not the focus of this work. The present work claims that besides animacy, the frequency, concreteness, and prototypicality of nominal referents should be involved as important non-discourse determinants of the topic of a sentence.

From what has been discussed so far, the following points may be made. First, a number of linguists have agreed that the primary psychological function of a grammatical subject is to refer to the cognitive point of departure, the starting point, or the psychological subject in the mind of a speaker. Second, the traditional notion of the speaker's cognitive point of departure has been characterized as a purely descriptive discourse relation within a sentence. As a result, the psychological aspects of the speaker's cognitive point of departure have generally been neglected, although the cognitive point of departure is by nature psychological. Third, a number of linguists seem to have failed to distinguish between the discourse and non-discourse determinants of the cognitive point of departure. It should be noted, however, that they generally suggest that there is an interaction between the discourse and the non-discourse determinants with regard to a speaker's choice of the cognitive point of departure and thus of the grammatical subject. For example, human referents may be more likely to be chosen as the topic of discourse than non-human ones; human topics are more likely to be chosen by speakers as grammatical subjects than non-human topics. Fourth, the discourse and non-discourse aspects of nominal referents should be viewed, to some extent, as being independent of each other with respect to the selection of grammatical subjects. It is then necessary to investigate separately the discourse and non-discourse determinants, the latter of which is the main focus of the present work (for the former, see Itagaki, 1982).

2.2.2 Psycholinguistic Studies of Cognitive Point of Departure

There have been a number of proposals characterizing the traditional notion of the speaker's cognitive point of departure as the primary psychological function of a grammatical subject. One popular way is in terms of the perceptual salience of nominal referents (e.g., Clark & Chase, 1974; Engelkamp & Zimmer, 1983; Flores d'Arcais, 1973; MacWhinney & Bates, 1978). They basically assume that if some nominal referents are perceptually more salient (e.g., size, shape, brightness, and so on) than others, they may be more available in the mind of a speaker. As a result, they are more likely to be chosen as cognitive points of departure and thus as grammatical subjects.

Clark and Chase (1974), for example, found that subjects preferred *The star is above/below the line* to *The line is below/above the star* to a significant degree, when they were asked to describe a picture in which one simple star figure was depicted either above or below one simple line. Note that the grammatical subject of the preferred sentence is *The star*, not *The line*, which suggests that certain geometric shapes are perceptually more salient than others.

Furthermore, Engelkamp and Zimmer (1983) attempted to determine whether or not the sizes of objects are indicative of their perceptual salience. They provided subjects with simple pictures in which one small object was depicted to either the right or the left of one large object, and asked the subjects to choose between two sentences describing the picture, one having the small object as its grammatical subject, the other the large one as its grammatical subject. It was found that subjects chose the sentences in which the grammatical subjects referred to the small objects 73% of the time. This can be taken to mean that the smaller objects are perceptually more salient than the larger ones, and are thus chosen by the subjects as their cognitive points of departure and as grammatical subjects.

Johnson-Laird (1968a, 1968b, 1977) provided experimental evidence seemingly contrary to the results of Engelkamp and Zimmer (1983). Johnson-Laird (1968a) gave subjects a pair of active and passive sentences, for example, *Red follows blue* and *Blue is followed by red*, and a slip of paper on which two narrow (5in. x 3/4in.) rectangles were drawn. The subjects were then asked to illustrate what they thought the two sentences meant by colouring the two rectangles with crayon. It was shown that the subjects of both the active and passive sentences were illustrated by larger coloured areas than the other noun phrases, and that the subjects of the passive sentences were illustrated by larger coloured areas than those of the active ones (cf. Costermans & Hupet, 1977). These results suggest that the larger coloured areas are perceptually more salient than the smaller ones, which does not seem to be consistent with what Engelkamp and Zimmer found. This implies that the notion of perceptual salience is by nature relative and hard to define.

Another similar way to characterize the traditional notion of the speaker's cognitive point of departure as a primary psychological function of grammatical subjects is in terms of a traditional Gestalt principle, called "figure" vs. "ground" (e.g., MacWhinney, 1977; Osgood, 1980; Osgood & Bock, 1977). It is generally assumed that some stimuli in a peceptual field are easier to perceive as "better organized wholes" (i.e., figures) than others, with the latter as grounds (Anderson, 1980; Glass, Holyoak, & Santa, 1979). From this, it follows that the former are more available in the mind of a speaker than the latter, with the former being likely to be chosen as the cognitive point of departure and thus as the grammatical subject.

Talmy (1978), for example, elaborates the notions of figure and ground as two semantic roles in motion or location predications. He defines the semantic roles, Figure and Ground, as follows:

> The Figure object is a moving or conceptually movable point whose path or site is conceived as a variable the particular value of which is the salience issue. (original emphasis, p. 627)
>
> The Ground object is reference-point, having a stationary setting within a reference-frame, with respect to which the Figure's path or site receives characterization. (p. 627)

For instance, in *John is beside Mary* and *Mary is beside John*, the referents of *John* and *Mary* function as the Figure and Ground and as the Ground and Figure, respectively. In both sentences, the grammatical subjects serve to specify the referents which function as the Figure, not Ground.

It may be misleading to discuss the notions of salience and figure as independent of each other with respect to the choice of the cognitive point of departure and subject selection (Osgood, 1980; Osgood & Bock, 1977). This is because the perceptual salience of stimuli may be among the so-called "gestalt principles," (e.g., "continuity," "proximity," "similarity," and so on) by which figures can be separated from grounds in perceptual fields. In other words, the difference between "well-organized" and "ill-organized" wholes may be attributed to these gestalt principles. Since the notion of salience is quite vague (Berlyne, 1960) and any precise definition of it is not found in the psychology literature, it seems better not to view it as one decisive gestalt principle. However, this does not mean that the notion of salience is not useful.

It is clear that the notions of salience and figure may be non-discourse determinants of the cognitive points of departure and thus of grammatical subjects. That is, the notions of salience and figure may be related to the notion of the cognitive availability of nominal referents in some way. For example, animate referents can be considered to be more salient than inanimate ones in the sense that people are generally more attentive to the former. Moreover, prototypical instances of a category seem to be easier to perceive as figures than others, since the former are generally "better" instances than the latter (see Section 2.1).

There have been several psycholinguistic studies whose purpose is to investigate directly the relationship between the inherent properties of nominal referents and their likelihood of being the cognitive point of departure (e.g., Clark, 1965; Clark & Begun, 1971; Dewart, 1979; Harris, 1978; James, 1972). In particular, most of these studies were focused on the choice between simple active and passive sentences.

Clark (1965) provided his subjects with the following types of simple incomplete active and passive sentences: "The ____ ____ed the ____" and "The ____ was ____ed by ____," and asked them to complete the sentences by guessing the missing words and writing them in the blanks. It turned out that the subjects of the active and passive sentences were animate nouns 81.5 % and 68.3 % of the time, respectively, while the second nouns of the active and passive sentences were animate only 26.7 % and 45.3 % of the time, respectively. This simply suggests that the grammatical subjects of active and passive sentences are more likely to be animate noun phrases than the second noun phrases, regardless of whether the sentences are active or passive.

Clark and Begun (1971) argue for the following hierarchy of grammatical subjects in simple English active sentences: Human > Animal > Concrete - Count (e.g., books and trees) > Concrete - Mass (e.g., grain and snow) > Abstract - Count (e.g., facts and weeks) > Abstract - Mass (e.g., harm and growth), where "A>B" means that A is more acceptable as a grammatical subject than B. They go on to hypothesize that nouns higher in the hierarchy could substitute for the grammatical subjects of active sentences, which are lower in the hierarchy, without losing the sensibleness of the sentences. For example, suppose that a human noun, *the man*, is substituted for the subjects of the following sentences: *The dog / The book / The snow / The fact / The growth surprised the people*. The resulting sentence, *The man surprised the people*, is semantically as acceptable as the originals. On the other hand, *The book, The snow, The fact*, and *The growth* could not substitute for the subject, *The man*, of the following: *The man watched the game*. That is, *The book / The snow / The fact / The growth watched the game* are semantically anomalous. (Note that *The dog watched the game* is semantically acceptable.)

Clark and Begun actually had subjects rate the sensibleness of simple English active sentences whose subjects were substituted with other nouns either higher or lower on the hierarchy. For example, the subject, *The snow*, of *The snow pressed the grass* may be replaced by *The man, The dog, The book, The fact, The harm*. The resulting sentences are *The man / The dog / The book / The fact / The harm pressed the grass*. They found that the subjects rated as higher the sensibleness of the first three sentences than that of the last two. (Note that the first three substitutes (i.e., *the man, the dog*, and *the book*) are higher in the hierarchy than *the snow*, while the last two substitutes (i.e, *the fact* and *the harm*) are lower in the hierarchy than *the snow*.) From this, Clark and Begun concluded that *the man, the dog*, and *the book* might be more acceptable as grammatical subjects than *the snow, the fact*, and *the harm*.

A few additional points must be made here. First, Clark and Begun are not concerned with any particular psychological function of a grammatical subject, but with the "psychological reality" of Chomsky's (1965) "selectional restrictions." This means that they view the selection of grammatical subjects as dependent only on the selectional restrictions of verbs. If this is true, speakers would have to choose the verbs prior to the grammatical subjects, and verb-initial languages would then be preferred to verb-final languages. However, this tendency does not seem to be true. Keenan (1978) suggests that the comprehension of some verbs is largely dependent on the the comprehension of the grammatical subjects. For example, the meanings of *is strong* in *Tom is strong* and *The chair is strong* are different, namely that Tom can exert a lot of force and The chair can stand a lot of force. Consequently, if the hearers are first given the meaning of the subject, either *Tom* or *The chair*, they would have little trouble in choosing one particular meaning of the predicate, *is strong*; if they first hear the predicate, they would have to wait to choose one particular meaning of the predicate until they hear the subject, that is, they would have more load in so-called "short-term" memory. From this, Keenan claims that verb-initial languages are cognitively less acceptable than verb-final ones.

Second, Clark and Begun failed to take into account other inherent properties of nominal referents which seem to be involved in the choice of grammatical subjects, for example frequency and prototypicality. Third, Clark and Begun focused on simple English active and passive sentences. There is no doubt that the subject accessibility hierarchy needs to be tested using different sentence types and experimental procedures.

From the point of view of developmental psychology, Dewart (1979) and Harris (1978) sought to determine whether there is any correlation between young children's use of passive sentences and the animacy of nominal referents. Dewart reported that when young children aged either six or eight years were asked to recall passive sentences in which the actors and the patients were animate and inanimate, respectively (for example, *The tractor was driven by the farmer*) they changed the voice at recall 59.38% of the time, namely *The farmer drove the tractor*. On the other hand, when they were asked to recall passive sentences whose actors and patients were inanimate and animate, respectively (for example, *The farmer was hit by the tractor*), they changed the voice at recall only 21.88% of the time. These results clearly suggest that young children tend to choose animate nouns rather than inanimate ones as grammatical subjects.

Harris (1978) asked subjects, both young children and adults, to describe simple pictures in which two participants engaged in certain activities (for example, A cat is patting a ball, A bicycle is pushing a rabbit, and so on). It is important to note that each pair of participants was chosen so as to be either the same or different in animacy. The young subjects were either five or ten years old, while the adult ones were 17 years old. It was found that when the actors and patients in the pictures were animate and inanimate, respectively, the younger children and adult subjects produced passive sentences, for example, *A ball is being patted by a cat*, only 0.3% and 0.5% of the time, respectively. When the actors and the patients were inanimate and animate, respectively, passive sentences, e.g., *A rabbit is being pushed by a bicycle*, were produced by the young children and the adult subjects 17.3% and 10.6% of the time, respectively. From this, Harris concluded that one of the functions of passive sentences might be to place animate before inanimate nouns, although the overall use of passive sentences was small in comparison with the overall use of active sentences (Weiner & Labov, 1983).

These studies imply that the animacy of nominal referents may have significant effects on the choice of grammatical subjects, regardless of whether speakers are young children or adults. However, it should hardly be surprising that this holds true for young children, since the thinking of young children tends to be more egocentric and animistic than that of adults (see Section 2.1). It is partly on the basis of this line of reasoning that the present work hypothesizes animacy as one of major non-discourse determinants of the cognitive point of departure and thus of the grammatical subject.

Concerning the imagery - concreteness of nominal referents, James (1972) and James, Thompson, and Baldwin (1973) made an interesting finding. James (1972) found that the recall of the subjects and objects in basic transitive sentences had to do with the imagery (i.e., image-value) / concreteness of their referents rather than their syntactic roles. This means that the higher the imagery of the referents are, the better they are recalled, whether they are the subjects or objects in the sentences to be recalled. It was also reported by James, Thompson, and Baldwin (1973) that high image-value object nouns of passive sentences tended to be recalled as the subject nouns of the corresponding active ones, with a change of voice at recall (also see James & Abrahamson, 1977). One critical point must be made: they failed to realize that there might be some interaction between the imagery (i.e., concreteness) and animacy of the referents. The effects of the animacy of nominal referents on the choice of voice at recall have been found by Dewart (1979) and Harris (1978) as discussed above.

Taken together, these studies suggest that some inherent non-discourse properties of nominal referents, especially animacy, should not be ignored as determinants of the cognitive point of departure and thus of subject selection. The basic hypothesis of this work proposes that a speaker's choice of cognitive point of departure as the primary psychological function of a grammatical subject is related to the cognitive availability of nominal referents.

Finally, it is necessary to discuss two other specific psycholinguistic characterizations of the traditional notion of the speaker's cognitive point of departure as the primary psychological function of a grammatical subject: Ertel's (1977) "a speaker's reference point of sentential construction" and Zubin's (1979) "a speaker's focus of interest."

Ertel proposes the notion of a speaker's reference point of sentential construction as a specific cognitive operation or process. He describes it as follows:

> It is assumed here that one of the basic mental operations underlying sentence construction is a certain manner of selection that may be called *nominal seizing*. The speaker seizes one and only one of cognitive units that offer themselves as nominal candidates within the realm of what is going to be uttered. The cognitive unit that has been seized is the primary reference point of the sentential construction. Once the reference point is decided upon, the rest of the sentence - the other nominal units included - will be set in relation to this point. Its main role is to serve as a kind of cognitive device for fixing the sentential construction. As a result, it will be represented linguistically as a noun phrase preceding the verb and nonsubject noun phrase. Apparently we are dealing here with the noun phrase that has traditionally been classified as the grammatical subject. (original emphasis, pp. 146-147)

It seems obvious that the notion of the speaker's reference point is one possible way to characterize the traditional notion of the speaker's cognitive point of departure.

The question is how a speaker chooses such a reference point, namely the seized nominal unit. Ertel goes on to claim that the psychological distance between the speaker himself and nominal referents in the event to be verbalized may be reflected in his choice of the reference point. In other words, the psychologically closer a nominal referent is to the speaker, the more likely it is to be seized as the reference point from which he can proceed to formulate the sentence. Obviously, it is possible to capture the notion of psychological distance between the speaker himself (i.e., ego) and nominal referents in terms of their cognitive availability for human cognition. To be more specific, such psychological distance may be related to the animacy of nominal referents, which is hypothesized as a major non-discourse determinant on the grounds of the assumption that human thinking is subject to egocentric or anthropocentric bias toward humans and related entities and against others.

Finally, Zubin (1979) attempts to characterize the traditional notion of the speaker's cognitive point of departure as the focus of interest of the speaker. He bases the notion of the speaker's focus of interest on two prominent psychological characteristics of human cognition, "selective attention" and "egocentrism." People tend to choose some particular entities as the focus of attention, since the resource of human attention is severely limited; they are also susceptible to egocentric bias towards themselves and against others, when they think about the internal and external world (Anderson, 1980; Glass, Holyoak, & Santa, 1979; Lachman et al. 1979).

Zubin assumes that a speaker's focus of interest is a function of both the selective attention and egocentric bias of the speaker, which together yield the following hierarchy of the speaker's foci of interest: speaker > hearer > other human (central > peripheral) > (inanimate) concrete > abstract - human related (e.g., thought, knowledge, etc.) > abstract, where "A>B" means that A is more likely to be chosen by a speaker as his focus of interest and thus as a grammatical subject than is B. On the basis of the analysis of a text sample, Zubin found that the author of the book himself, inanimate concrete entities, and abstract entities were realized by grammatical subjects about 75%, 40%, and 25% of the time, respectively. This is consistent with the above hypothetical hierarchy, that is, speaker > inanimate concrete > abstract.

It is possible, however, to argue that a speaker's choice of the focus of interest as the primary psychological function of a grammatical subject may be determined by more than the animacy - concreteness of nominal referents. For example, the frequency and prototypicality can be considered to be possible factors. It is also of importance to see whether or not there is any interaction among these inherent properties with regard to the choice of the cognitive point of departure in the mind of a speaker as the primary psychological function of a grammatical subject.

What has been discussed so far may be summarized on the following four points. First, a number of psycholinguists agree that the primary psychological function of a grammatical subject is to refer to the cognitive point of departure in the mind of a speaker. Second, they have endeavoured to characterize the traditional notion of the speaker's cognitive point of departure in terms of certain psychological factors, for example the notions of perceptual salience and figure, or as certain cognitive strategies, for example Ertel's "a speaker's reference point of sentential construction" and Zubin's "a speaker's focus of interest." Third, although various inherent properties of nominal referents have been proposed as non-discourse determinants of a speaker's cognitive point of departure, they seem to be captured by the notion of cognitive availability of the referents, as defined in this work. Fourth, it is crucial to investigate the relationship between the cognitive availability of nominal referents and their likelihood of being chosen by a speaker as his cognitive point of departure and thus as the grammatical subject.

2.3 Some Implications for Linguistics

The present work offers two basic psycholinguistic implications for linguistics. One is to make a psycholinguistic contribution to the clarification of the notion of grammatical subject. The other is to seek to provide a psycholinguistic account of the nearly universal relative order of subjects before objects (Greenberg, 1966).

Keenan (1976) made an interesting attempt to establish a "universal" definition of subjects in basic sentences (cf. Johnson, 1977a, 1977b). As his basic approach, he states:

> We are not free to define a notion like "subject" in any way that suits our purposes. There is a large body of lore concerning the notion, and any proposed definition must at least largely agree with the traditional, and to some extent, pretheoretical usage of the term. Our approach then will be to collect a large and diverse set of cases from different Ls [Languages] in which our pretheoretical judgements of subjecthood are clear. Then we shall attempt to abstract from this set a set of properties which are characteistic of subject NPs and then try to determine some combination of the characteristic properties which will be jointly necessary and sufficient to pick out the subject of an arbitrary sentence in an arbitrary L in a way that is in conformity, of course, with our pretheoretical intuitions in the clear cases. ... Note further, that on this type of definition, subjects of certain sentences, and more generally of certain sentence types, will be more subject-like than the subjects of others. The reason is that they will exhibit more of the complement of properties which

> characterize b-subjects [subjects in basic sentences] in general. Thus the subjecthood of an NP (in a sentence) is a matter of degree. (pp. 306-307)

The important point is that some of the properties of subjects have to do with the inherent properties of nominal referents, which are the focus of the present work.

Keenan (1976) then identifies the four major categories of the characteristic properties of the subjects in basic sentences: "autonomy properties," "case marking properties," "semantic role," and "immediate dominance." The two conceptual autonomy properties of "independent existence" and "autonomous reference" are of relevance to the present thesis. It seems that these two properties are compatible with the cognitive availability of nominal referents as defined in this work.

What is meant by independent existence is that the existence of the referents of subjects is less dependent on the action expressed by verbs than that of the referents of objects. For example, the subject, *a student*, of the following sentences can be considered to exist more independently of the actions expressed by the verbs than the objects: *A student came up with a good idea, wrote a short essay, defined a new mathematical term*. It is important to note that the existence of the referents of *a good idea, a short essay*, and *a new mathematical term* is simply the result of the actions of *coming up with, writing*, and *defining*, respectively.

It seems possible to argue that the dependent or independent existence of nominal referents is related to their animacy and/or concreteness. That is, the more animate and/or concrete the nominal referents are, the more likely they are to exist independently of any action or activity. Returning to the above sentences, the referents of the objects, *a good idea, a short essay*, and *a new mathematical term*, are obviously less animate and concrete than that of the subject, *a student* (see Section 2.1).

In discussing autonomous reference, Keenan points out two properties. First, he argues that subjects are always more highly referential than objects or other nominal constituents. Accordingly, highly referential nouns, such as personal pronouns, proper nouns, and demonstratives, may occur in the subject positions in sentences. Second, it is maintained that subjects tend to refer to more absolute referents than objects or other nominal constituents. In other words, if a sentence is true, the existence of the referent of the subject is required, while the existence of the referent of the object is not necessarily required. For example, the referents of the subjects in the following sentences, *A student talked about an ideal government* and *A boy asked Santa Claus for a present*, are required to exist, if the sentences are true. On the other hand, the referents of the objects need not exist. That is, the referents of the subjects can be viewed as being more absolute than those of the objects.

It seems possible to argue that these two autonomous properties of subjects may be captured in terms of the animacy and/or concreteness of their referents. Animate and/or concrete nominal referents are, generally, absolute and/or highly referential, while inanimate and/or abstract nominal referents are not necessarily absolute and highly referential (see Section 2.1).

From these considerations, it follows that these two conceptual properties of subjects, independent existence and autonomous reference, may be consistent with the two hypotheses of the present work: the Concreteness and Animacy Hypotheses. The Concreteness Hypothesis states that the more concrete a nominal referent is, the more likely it is to be chosen by a speaker as his cognitive point of departure and as a grammatical subject in a sentence. According to the Animacy Hypothesis, the more animate a nominal referent is, the more likely it is to be chosen as the cognitive point of departure and therefore as the grammatical subject.

It is important, therefore, for the further understanding of the universal aspects of subjects to investigate the cognitive availability of nominal referents such that subject selection can be explained from a psycholinguistic point of view. It is also necessary to investigate such universal aspects of subjects in an empirical manner.

The present work may have psycholinguistic implications for the nearly universal tendency for grammatical subjects to precede other nominal constituents. (Throughout this thesis, this tendency will be referred to as "SO order," where "O" indicates any non-subject nominal constituent.) It is well known that nominal subject noun phrases in so-called "basic" or "unmarked" sentences tend to precede object or other noun phrases. This is explicitly stated by Greenberg (1966):

> Logically, there are six possible orders: SVO, SOV, VSO, VOS, OSV, and OVS. Of these six, however, only three normally occur as dominant orders. The three which *do not occur at all or at least are excessively rare*, are VOS, OSV, and OVS. These all have in common that the object precedes the subject. This gives us our first universal: *Universal 1*. In declarative sentences with nominal subject and object, the dominant order is almost always one in which the subject precedes the object. (original emphasis, pp. 76-77)

This dominant SO order is also claimed by a number of other linguists (e.g., Dik, 1978, 1980, 1983; Pullum, 1977; Steele, 1978, Ultan, 1969; Vincent, 1979).

Ultan, for example, reports on the following percentages of the four basic word order types of a sample of 79 randomly chosen languages: SOV - 44.0%, SVO - 34.6%, VSO - 18.7%, and VOS - 2.7%. The point is that neither OSV nor OVS languages were found in the large sample; and that, as far as basic or unmarked sentences are concerned, the subjects were shown to precede the objects in more than 97 percent of the sample of languages. It is in this empirical sense, not in a rationalistic sense, that the SO order is argued to be dominant (Comrie, 1981).

A psycholinguistic account for the nearly universal SO order seems to be composed of two basic steps. The first has to do with the psychological function a grammatical subject serves. (The present work assumes that the primary function is to refer to the cognitive point of departure in the mind of a speaker.) The second has to do with the type of psychological strategy responsible for the nearly universal tendency for grammatical subjects to precede other nominal constituents. Such strategies may be related to the basic determinants of the cognitive point of departure as the primary psychological function of a grammatical subject.

From a linguistic point of view, Tomlin (1979) proposes three basic principles by means of which the dominant SO order may be explained. He calls two of them "expected first principle" and "animated first principle." As the expected first principle, he states: "This principle is an attempt to make more precise and explicit the idea that 'old' information precedes 'new' information" (p. 13). The animated first principle states: "... in basic transitive clauses the NP which is most 'animated' will precede other NPs" (p. 14). As mentioned earlier, Tomlin defines the animatedness hierarchy in terms of both conceptual animacy and semantic case hierarchies.

The point is that these two principles are clearly consistent with the two types of determinants of the cognitive point of departure and thus of the grammatical subject. It seems reasonable, therefore, to seek to explain these two accounts for the dominant SO order from a psycholinguistic point of view.

As stated earlier, given-new information has been extensively discussed and investigated in the psycholinguistic literature as a typical discourse determinant of the cognitive point of departure in the mind of a speaker and thus of the selection of grammatical subjects (e.g., Bock, 1977; Bock & Irwin, 1980; Engelkamp & Zimmer, 1983; Clark & Haviland, 1977; Haviland & Clark, 1974). They generally found that subjects tended to choose given information as grammatical subjects rather than new information. More importantly, it has been shown that given information tends to be placed in an earlier surface position in a sentence than new information.

Those authors also base the given-new order on the following basic cognitive strategy for sentence comprehension: in order to understand a sentence, a hearer tends to execute the following three basic steps: Step 1 - Identification of Given Information in Memory, Step 2 - Identification of New Information in the Sentence, and Step 3 - Integration of New to Given Information in Memory (Clark & Haviland, 1977, Haviland & Clark, 1974; Hornby, 1974). It is clear that if hearers follow basically these three steps, the given-new word order may help them to comprehend the sentences by proceeding through the first step, through the second step, to the third step. In addition, Bock (1977) and Bock and Irwin (1980) showed that their subjects tended to follow the given-new word order, when they were asked to recall and to produce English sentences (also see Osgood & Bock, 1977).

Clark and Clark (1978) base their psycholinguistic account for the dominant SO order on this line of reasoning. They state:

> Note that the subject (S) comes first in the two commonest language types, SVO and SOV (about 80 percent of Ultan's sample), and it precedes the object in all but the rarest type, VOS (less than 2 percent of Ultan's sample). Why subjects are placed early seems fairly clear. *People tend to express given information, what is already known, to the listener, before new information, what is not already known.* This tendency appears to be universal. (emphasis added, p. 258)

It is important to note that the nearly universal SO order may be explained as reflecting the given-new strategy for sentence comprehension and production.

The present work is concerned with a second account for the dominant SO order, i.e., non-discourse determinants of the cognitive point of departure and Tomlin's (1979) "animated first principle." What this work seeks to demonstrate is that nominal referents of high cognitive availability are more likely to be chosen by speakers as grammatical subjects. Animacy is among the possible inherent properties of nominal referents which may be hypothesized as major parameters of the cognitive availability of the referents. In addition, this work has hypothesized that the frequency, concreteness, and prototypicality may contribute to their availability in human cognition. The general hypothesis to be tested is in turn that the more frequent, concrete, animate, and/or prototypical a nominal referent is, the more likely it is to be a speaker's cognitive point of departure and thus a grammatical subject.

The next question is how to relate these non-discourse determinants to some psychological strategy which may be responsible for the sentence-initial positions of grammatical subjects. The answer seems to rest on Bock's (1982) and Bock and Irwin's (1980) hypothesis that some nominal referents are easier to lexicalize than are others, with the former occuring in earlier surface positions than the latter. That is, it may be possible to explain the nearly universal relative order of subjects before other nominal constituents in terms of their ease of lexicalization.

In order to pursue this line of accounting for the dominant SO order, what needs to be done first is to test the hypothesis that grammatical subjects tend to refer to nominal referents of high cognitive availability rather than to those of low cognitive availability. If this is true, grammatical subjects can be considered to be easier to lexicalize than others, with the former being likely to occur in earlier surface positions in sentences than the latter.

Bock (1982), indeed, states the following in connection with this point:

> For example, several of the theories of constituent ordering cited above assume that humans process information egocentrically and are, therefore, predisposed to attend to personally relevant stimuli. Among such personally relevant stimuli are other animate beings, particularly human animate beings. Thus, animate entities should tend to occur early in sentences more often than inanimate entities. [This is hypothetically because the animate entities are easier to lexicalize than the inanimate ones.] ... there is a well-known universal preference for subject-object ordering in the basic constituent orders of the world's languages (Greenberg, 1966; Pullum, 1977). *This preference is presumably correlated with agentivity/animateness in the subject and nonagentivity/inanimateness in the object*. (emphasis added, p. 15)

From this, it follows that the association between the animacy of nominal referents and their likelihood of being grammatical subjects must first be psycholinguistically verified. In fact, not only animacy, but also the other non-discourse properties must be investigated as such (cf. Comrie, 1981; Tomlin, 1979).

Finally, it is necessary to point out that Dik (1978, 1980) proposes to account for the dominant SO order in terms of a speaker's perpective or orientation from which he presents a state of affairs. He states:

> We interpreted Subj and Obj assignment by predication as a method of defining a perspective on the state of affairs designated by predication, starting with one of entities involved (the Subj), through a second entity involved. It stands to

> reason, then, that this perspective is mapped onto the linear structure of the sentence in the order Subj - Obj - Other rather than in some order. This implies that the minority of languages alleged to have dominant orders with the Obj before the Subj require special explanation. (p. 176)

This explantion for the SO order consists of the following two basic steps of research. First, the strong association between subjects and speakers' perspective must be verified linguistically and psycholinguistically (Itagaki & Prideaux, 1983). Second, the tendency for the subjects functioning as the speakers' perspective to precede other nominal constituents must be confirmed linguistically and psycholinguistically.

This thesis addresses the following critical question as part of a psycholinguistic clarification of the first point: Why are some entities likely to be selected as the targets of perspective rather than others? In other words, what kinds of non-discourse as well as discourse properties of nominal referents are correlated with the speaker's decision as to which perspective to take? This work assumes that such a decision is correlated with the following four inherent nominal properties: frequency, concreteness, animacy, and prototypicality. That is, the more frequent, concrete, animate and/or prototypical a nominal referent is, the more likely it is to be the target of a speaker's perspective (cf. Kuno & Kaburaki, 1977; Itagaki, 1982).

3. EXPERIMENTS

The present thesis is focused on the four inherent non-discourse properties of nominal referents as the major determinants of the choice of grammatical subjects. The general hypothesis has been developed that the more frequent, concrete, animate, and/or prototypical a nominal referent is, the more likely it is to be realized as a grammatical subject.

The basic hypothesis is divided into four specific hypotheses; namely, the Frequency, Concreteness, Animacy, and Prototypicality Hypotheses. A further purpose of the present work is to see whether there is any interaction among these hypotheses with regard to a speaker's choice of the cognitive point of departure as the primary psychological function of a grammatical subject. This constitues the fifth hypothesis, namely the Interaction Hypothesis.

In this work, three experiments were designed and carried out. The first experiment was designed to test the Frequency, Concreteness, Animacy, and Interaction Hypotheses. The second and third experiments were designed to test the Prototypicality Hypothesis.

3.1 *Experiment 1*

3.1.1 *Purpose*

The purpose of Experiment 1 is to test the Frequency, Concreteness, Animacy, and Interaction Hypotheses. These are:

> *The Frequency Hypothesis.* The more frequent a nominal referent is, the more likely it is to be chosen by a speaker as his cognitive point of departure and thus as the grammatical subject in the sentence.
>
> *The Concreteness Hypothesis.* The more concrete a nominal referent is, the more likely it is to be chosen by a speaker as his cognitive point of departure and thus as the grammatical subject in the sentence.
>
> *The Animacy Hypothesis.* The more animate a nominal referent is, the more likely it is to be chosen by a speaker as his cognitive point of departure and thus as the grammatical subject in the sentence.
>
> *The Interaction Hypothesis.* There is some interaction among the inherent properties of nominal referents with regard to a speaker's choice of cognitive point of departure and thus of the grammatical subject in the sentence.

3.1.2 *Method*

Subjects. Thirty-one University of Alberta students, 17 females and 14 males, participated in Experiment 1 as volunteers. Nine were graduate students; all others were undergraduates. All the subjects were native speakers of

North American English.

Materials. In Experiment 1, the subjects were provided with a list of nouns, and then asked to compose either a single sentence or a short passage using each of the given nouns.

The target nouns were selected so as to differ from each other in at least one of the following properties: frequency of usage (high vs. low), concreteness (concrete vs. abstract), and animacy (human, animate non-human, vs. inanimate). As a result, the nouns fell into one of the following eight classes: High Frequency Human (HHUM), e.g. *a boy* and *a professor*, High Frequency Animate Non-Human (HANH), e.g. *a dog* and *a cow*, High Frequency Object (HOBJ), e.g. *a car* and *a gun*, High Frequency Abstract (HABS), e.g. *a fact* and *freedom*, Low Frequency Human (LHUM), e.g. *a mentor* and *an idler*, Low Frequency Animate Non-Human (LANH), e.g. *a wombat* and *a nighthawk*, Low Frequency Object (LOBJ), e.g. *rickshaw* and *wampum*, Low Frequency Abstract (LABS), e.g. *affluence* and *technocracy*. Throughout this thesis, these acronyms will be used to refer to the corresponding noun classes. In the experiment, no articles were used with the nouns.

For each of these eight classes, 10 nouns were chosen. The nouns were selected according to the following procedures. First, two graduate students (native English speakers) at the Department of Linguistics, University of

Alberta, chose about 15 nouns for each of the eight classes. They were simply asked to list nouns which they thought belonged to each class and which would not be difficult for average university students to compose either a single sentence or a short passage about. Second, two independent frequency counts were obtained for each of these nouns, one from Thorndike & Lorge (1944) and the other from Carroll, Davies, and Richman (1971). Third, from each of the eight classes, 10 nouns were chosen; they were either the 10 highest or lowest frequency count words of the nouns selected by the two graduate students. As a result, 80 nouns were chosen for Experiment 1. These words and their frequency counts are shown in Appendix A.

Two types of booklets were then designed, one for the "single sentence" task and the other for the "passage" task; 20 "single sentence" and 20 "passage" booklets were prepared. (The size of each page of the booklets was 8in. X 11in.) The "single sentence" booklets were composed of one instruction page and 20 other pages, with four nouns typed from top to bottom on the left side of each page. Similarly, the "passage" booklets had one instruction page and 40 other pages, with two nouns typed on the left side of each page, one at the top and the other in the middle. The orders of the typed nouns and pages were randomized.

Procedure. As mentioned earlier, there were two types of tasks. One was for the subjects to compose a single sentence using each of the 80 given nouns, namely the "single sentence" task. In the other task, the subjects were simply asked to compose a short passage using each of the 80 given nouns. As explained above, the 80 nouns consisted of 10 nouns of each of the eight classes. According to the basic hypothesis of the present work, these eight classes of nouns may differ in the frequencies with which they are realized as the grammatical subjects in the clauses in either the single sentences or the passages to be produced.

The two types of tasks, "single sentence" and "passage," were designed in order to test the the basic hypothesis in two different situations. In other words, it was of interest to see whether the difference between these two types of tasks might cause any difference in the relationship between the inherent non-discourse properties of nominal referents and their likelihood of being grammatical subjects.

In the instructions, the following points were emphasized. First, the 80 nouns were independent of each other and thus the subjects were supposed to ignore all other nouns, when they were focusing on one particular noun. Second, there was no restriction on the content, the form or style, the length, the complexity, and etc. of a sentence, as long as it was a complete, single sentence. Nor was there any restriction on the content, the form, or the length (within the half an 8in. X 11in. page) of a passage. Third, the subjects were unrestricted as to how to use each given noun, for example plural or singular forms, compound forms, positions in sentences, and so on. The subjects were encouraged to write whatever they wished about each given noun; they were also allowed to skip some nouns, if they felt that they could not formulate single sentences or passages with them.

Forty subjects were provided with either the "single sentence" or "passage" booklets; the assignment of booklets was made at random. They were allowed to do the task at their convenience. They were asked to bring the booklets back as soon as they finished. On the average, the "single sentence" and "passage" tasks took 70 minutes and three hours, respectively, to complete. In the end, 17 "single sentence" and 14 "passage" booklets were completed and returned.

3.1.3 Analyses and Results

In order to follow the analyses and results discussed below, the reader is urged to consult first the sample single sentences and passages produced by the subjects. For this purpose, some samples of the single sentences and passages are presented in Appendices B and C, respectively.

The analysis of the data focused on the frequency with which each given noun was realized as the grammatical subject, more specifically the head noun of the subject noun phrase, and as other grammatical roles (e.g., direct and indirect object, complement, noun phrase in adverbial phrases, etc.). The grammatical subjects of both independent and dependent clauses were counted; the dependent clauses were ones which had subordinate conjunctions (e.g., *when*, *before*, *so that*, *as long as*, *now that*, and so on).

The following criteria were imposed in order to determine whether some data should be included or not. First, given nouns should not elicit particular verbal associates, particularly verbs. The reason for this was that it was necessary to keep the subjects from producing certain "idiosyncratic" or "idiomatic" sentences for some nouns, which could not be viewed as representing the likelihood of being grammatical subjects of the nouns. However, this point was not implemented as part of the instructions. This is because if this point was specifically instructed to the subjects, they would hesitate to write whatever they wished. As a result, one noun *writer* was excluded in all cases, since in the "single" task, it elicited a copula verb, *be* or *become*, almost dominantly, e.g., *Tom is a good writer* and *I am a writer*. Second, nouns were excluded from further analysis when they were used metaphorically, for example for the noun "sheep," one subject produced *If someone calls you a sheep, it is not compliment*, in which the noun *sheep* here is used in reference to a human. Third, nouns were excluded when their meanings were clearly misunderstood, misused, or confused with other nouns. For example, nouns used as verbs were excluded, as were nouns used "linguistically," as in the case of the item "idler," *The word, "idler," refers to the one who is lazy*. After some data were eliminated on the basis of these criteria, two frequencies with which the given noun was realized as grammatical subjects and other grammatical roles in the clauses were obtained for each single sentence and passage.

On the basis of these frequency counts, one ratio measurement was calculated as a primary measurement for each noun and each subject, for a total of 80 ratio measurements for each subject. The formula for the ratios is the following: $F(S)/F(S)+F(O)$, where "$F(S)$" and "$F(O)$" indicate the frequency counts of the given noun as grammatical subjects and as other grammatical roles, respectively. For example, suppose that for a given noun "professor," a subject produces a single sentence, *We don't like the professor, since he is terrible at teaching* or *Professors always complain that they are not well paid*. The ratios for these are $1/1+1 = 0.5$ and $1/1+0 = 1.0$, respectively. It is important to notice that these ratios represent the fact that in the first sentence, the test noun is used as both the subject and the direct object, while in the second sentence, it is used twice as the subject. Therefore, high ratios for some nouns mean that they are not only more likely to be grammatical subjects than to be in any other grammatical role, but also more likely to be grammatical subjects than other nouns. The same scoring method was used for the data from the passages in order to obtain the ratio measurement for each noun.

The mean ratio and standard deviation for each of the eight noun classes are shown in Table 1. For the sake of further inspection, the mean ratio for each of the 80 nouns is presented in Appendix D

In addition, for the sake of statistical analysis, these ratio measurements were transformed according to Ferguson (1976). (The formula of the transformation is ARCSIN(SQRT(X)), where "X" is a ratio.) An analysis of variance was then performed with the transformed ratio measurements in order to test the hypotheses. The between-subject independent variable was Task (Single Sentence vs. Passage); two within-subject independent variables were Frequency (High vs. Low), and Animacy - Concreteness (Human, Animate Non-Human, Object, vs. Abstract). The definitions of these variables and their conditions are as discussed above.

One main effect, Animacy - Concreteness, was significant ($F(3,87) = 7.447$, $p < 0.01$), and more interestingly, the interaction between Frequency and Animacy

Table 1

Mean Ratios and Standard Deviations

		HIGH FREQUENCY[1]			
Task	Stats	HUM	ANH	OBJ	ABS
Single	Mean	0.627	0.614	0.448	0.410
	SD	0.218	0.160	0.209	0.145
Passage	Mean	0.636	0.597	0.493	0.392
	SD	0.128	0.178	0.205	0.171

		LOW FREQUENCY[1]			
Task	Stats	HUM	ANH	OBJ	ABS
Single	Mean	0.440	0.640	0.394	0.476
	SD	0.210	0.229	0.174	0.175
Passage	Mean	0.474	0.587	0.396	0.436
	SD	0.239	0.195	0.212	0.203

$0.0 \leq \text{Score} \leq 1.00$

[1] HUM - Human, ANH - Animate Non-Human, OBJ - Object, and ABS - Abstract.

- Concreteness was also significant ($F(3,87) = 6.325$, $p < 0.01$). The detailed results are shown in Table 2 and Figure 1. The significant interaction clearly indicates that the main effect of Animacy - Concreteness applies only in the case of high frequency nouns (see Figure 1).

In order to test further the Frequency, Concreteness, and Animacy Hypotheses, a series of Newman-Keuls *a posteriori* tests were carried out with the mean differences between 28 possible pairs of the following eight conditions: the Human, Animate Non-Human, Object, and Abstract cases of the High and Low conditions. The results of the tests which are relevant to the above hypotheses are shown in Tables 3 and 4. (Furthermore, the detailed studentized ranges, Q values, and their significances are presented in Appendix E.)

The test of the Frequency Hypothesis was focused on the mean differences between the following four pairs: the Human cases of the High and Low Frequency conditions, the Animate Non-Human cases of the High and Low Frequency conditions, the Object cases of the High and Low Frequency conditions, and the Abstract cases of the High and Low Frequency conditions. As shown in Table 3, only the difference between the Human cases of the High and Low Frequency conditions turned out to be significant ($Q4(87) = 5.901$, $p < 0.01$).

Table 2

Results of Analysis of Variance (Ratios)

Source[1]	SS	df	MS	F	p
A	0.057	1	0.057	0.189	
S/A	8.788	29	0.303		
B	0.231	1	0.231	2.468	
AB	0.041	1	0.041	0.435	
BS/A	2.720	29	0.094		
C	3.773	3	1.258	17.447	**
AC	0.220	3	0.073	1.018	
CS/A	6.271	87	0.072		
BC	1.199	3	0.400	6.325	**
ABC	0.137	3	0.046	0.723	
BC/A	5.497	87	0.063		

[1] S - Subjects
A - Task (Single Sentence vs. Passage)
B - Frequency (High vs. Low)
C - Animacy-Concreteness (Human, Animate Non-Human, Object vs. Abstract)

$^{**}p<0.01$

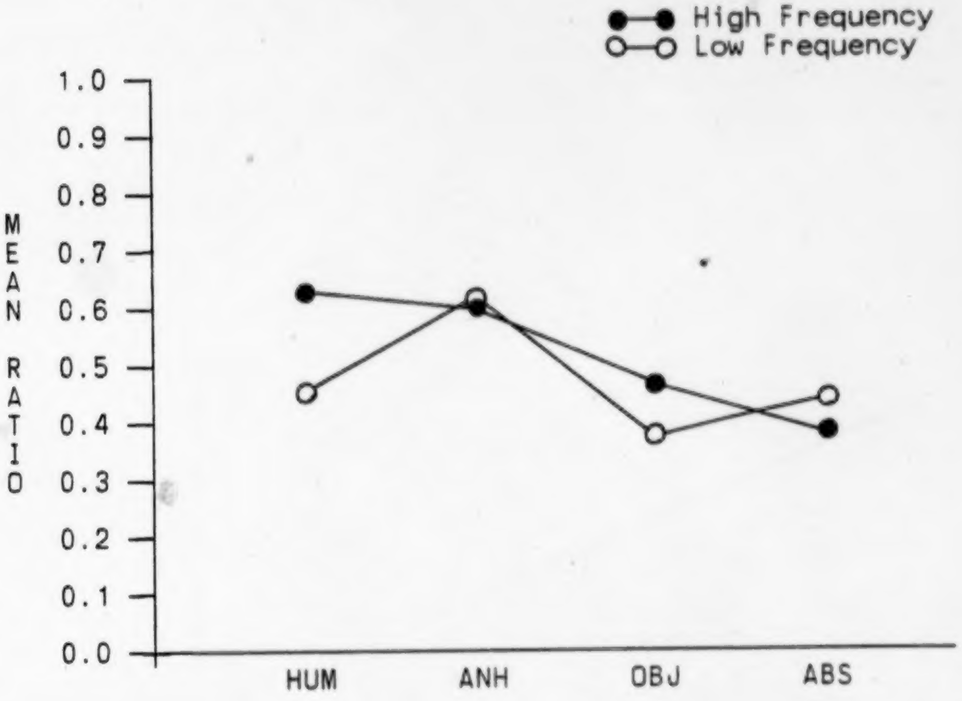


Figure 1. Interaction between Animacy-Concreteness and Frequency (Hum - Human, ANH - Animate Non-Human, OBJ - Object, and ABS - Abstract)

Table 3

Results of *A Posteriori* Tests (Ratios)

Pairs	Sig.
HHUM - LHUM	***
HANH - LANH	
HOBJ - LOBJ	
HABS - LABS	

***$p<0.01$

Table 4

Results of *A Posteriori* Tests (Ratios)

	HHUM	HANH	HOBJ	HABS		LHUM	LANH	LOBJ	LABS
HHUM	—		***	***	LHUM	—	***		
HANH		—	***	***	LANH		—	***	***
HOBJ			—		LOBJ			—	
HABS				—	LABS				—

***$p<0.01$

The mean difference between the Object cases did not reach significance level, although the mean ratios for the Object of the High Frequency condition (i.e., 0.448 and 0.493 in the single sentence and passage tasks) turned out to be slightly higher than those for the Object of the Low Frequency condition (i.e., 0.394 and 0.396 in the single and passage tasks). Furthermore, the overall results for the Animate Non-Human and Abstract conditions were found to be contrary to the expectation of the Frequency Hypothesis. That is, the mean ratios for the Animate Non-Human and Abstract cases of the Low Frequency conditions were slightly higher than those for the Animate Non-Human and Abstract cases of the High Frequency conditions, but the mean differences between these groups did not reach a significant level (see Figure 1 and Table 3). From these results, it follows that the Frequency Hypothesis can be confirmed with regard to human nouns, but not with regard to others.

The independent test of the Concreteness Hypothesis focused on the mean differences between the Object and Abstract conditions. These mean differences were not found to be significant in the case of either the High or the Low Frequency condition (see Table 4). This result implies that as far as the ratio measurements are concerned, the Concreteness Hypothesis may not be supported independently of the Animacy Hypothesis.

The Animacy Hypothesis was independently investigated on the basis of the tests of the mean difference between the following three pairs: the Human and Animate Non-Human conditions, the Human and Object conditions, and the Animate Non-Human and Object conditions. The results of the tests are shown in Table 4.

In the case of the High Frequency condition, the following two mean differences turned out to be significant: $Q3(87) = 5.501$, $p < 0.01$ for the Human and Object conditions and $Q2(87) = 4.414$, $p < 0.01$ for the Animate Non-Human and Object conditions. The mean difference between the Human and the Animate Non-Human conditions was not shown to be significant, in spite of the fact that the mean ratios for the Human (i.e., 0.627 and 0.636 in the single sentence and passage tasks) were slightly higher than those for the Animate Non-Human conditions (i.e., 0.614 and 0.597 in the single sentence and passage tasks). These findings suggest that there should be a highly significant clustering of human and animate non-human nouns on the one hand and object and abstract nouns on the other, namely animate versus inanimate (see Figure 1 and Table 4), and that this clustering holds only in the case of high frequency nouns.

Moreover, as expected, the mean difference between the Human and Abstract cases of the High Frequency condition was found to be highly significant ($Q6(87) = 7.631$, $p < 0.01$). It may follow that there is an additive interaction effect between the animacy and concreteness of nouns on the choice of grammatical subjects (see Figure 1).

Concerning the Low Frequency condition, the overall results of the Newman-Keuls *a posteriori* tests turned out to be contrary to both the Concreteness and Animacy Hypotheses (see Table 4). It was shown that the mean differences between the Human and Animate Non-Human conditions and between the Animate Non-Human and Object conditions were significant ($Q5(87) = 6.300$, $p < 0.01$ and $Q8(87) = 8.429$, $p < 0.01$). These results suggest that as far as low frequency nouns are concerned, animate non-human nouns were much more likely to be grammatical subjects than others, which is contrary to the Animacy Hypothesis. The significant interaction between Frequency and Animacy - Concreteness means that the effect of animacy and concreteness on the choice of grammatical subjects is limited to the case of high frequency nouns (see Figure 1).

Moreover, it turned out that the mean ratio for the Object case of the Low Frequency condition (i.e., 0.395) was slightly lower than that for the Abstract case of the Low Frequency (i.e., 0.456). This result is not, of course, predicted by the Concreteness Hypothesis. However, the mean difference between these two groups was not shown to be significant (see Figure 1 and Table 4).

In order to confirm these results further, another measurement was obtained for each subject. The measurements were calculated on the basis of the two different frequencies with which given nouns and other nouns (e.g., personal pronouns, proper nouns, etc.) were realized as grammatical subjects in dependent and independent clauses. The percentages of the frequencies of the 10 given nouns of each group to total frequencies of the 10 nouns and others as grammatical subjects were then calculated for each subject, for a total of eight percentage measurements for each subject. The mean percentage and standard deviation for each of the eight noun groups are presented in Table 5. These percentages were also transformed according to Ferguson (1976), using the same transformation formula as that used for the ratio measurements. (The transformation is based on the following formula: ARCSIN(SQRT(X)), where "X" is a percentage.)

With these transformed percentage data the same analysis of variance as with the transformed ratio measurements was performed. Similar results were obtained as expected. Three main effects of Task, Frequency, and Animacy - Concreteness turned out to be significant ($F(1,29) = 11.083$, $p < 0.01$, $F(1,29) = 8.176$, $p < 0.01$, and $F(3,87) = 26.384$, $p < 0.01$, respectively). The interaction effect between Frequency and Animacy - Concreteness was also significant ($F(3,87) = 6.929$, $p < 0.01$). The detailed results are shown in Table 6 and Figures 2 and 3. As can be seen in Figure 3, it was also confirmed that the main effect of Animacy - Concreteness was found to be significant only

Table 5

Mean Percentages and Standard Deviations

Task	Stats	HIGH FREQUENCY[1] HUM	ANH	OBJ	ABS
Single	Mean	57.64	55.19	39.88	31.43
	SD	20.15	17.51	18.28	14.41
Passage	Mean	44.55	43.73	29.31	19.41
	SD	12.63	14.58	15.31	12.26

Task	Stats	LOW FREQUENCY[1] HUM	ANH	OBJ	ABS
Single	Mean	38.13	55.86	33.82	40.89
	SD	18.91	22.64	16.16	17.89
Passage	Mean	30.51	37.61	22.28	18.38
	SD	20.35	12.41	13.48	8.98

$0.0 \leq \text{Score} \leq 100.00$

[1] HUM - Human, ANH - Animate Non-Human, OBJ - Object, and ABS - Abstract.

Table 6

Results of Analysis of Variance (Percentages)

Source[1]	SS	df	MS	F	p
A	2.966	1	2.966	11.083	**
S/A	7.761	29	0.268		
B	0.438	1	0.438	8.176	**
AB	0.076	1	0.076	1.412	
BS/A	1.554	29	0.054		
C	4.615	3	1.538	26.384	**
AC	0.192	3	0.064	1.098	
CS/A	5.073	87	0.058		
BC	0.956	3	0.319	6.929	**
ABC	0.115	3	0.038	0.830	
BCS/A	4.003	87	0.046		

[1] S - Subjects
A - Task (Single Sentence vs. Passage)
B - Frequency (High vs. Low)
C - Animacy-Concreteness (Human, Animate Non-Human, Object vs. Abstract)

**$p<0.01$

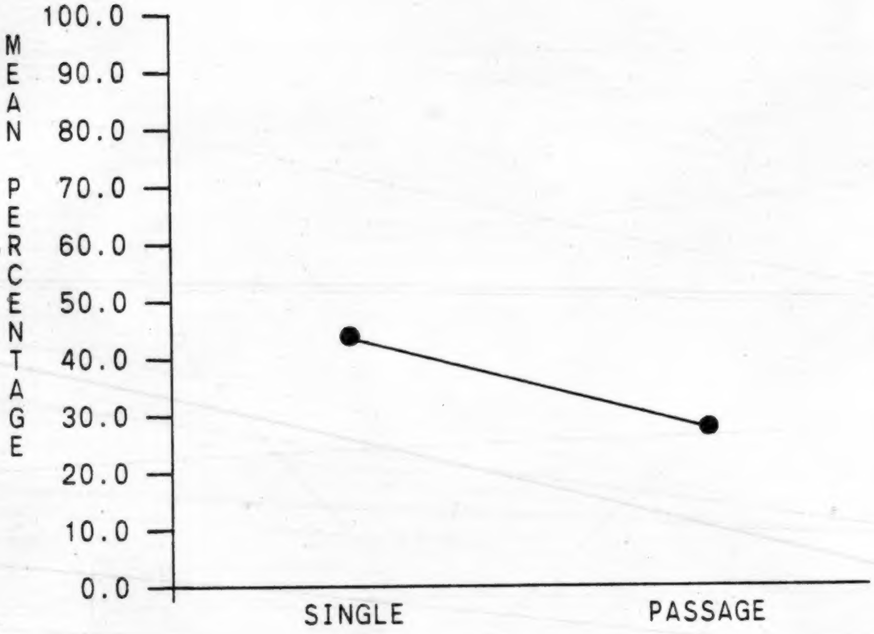


Figure 2. Main Effect of Task

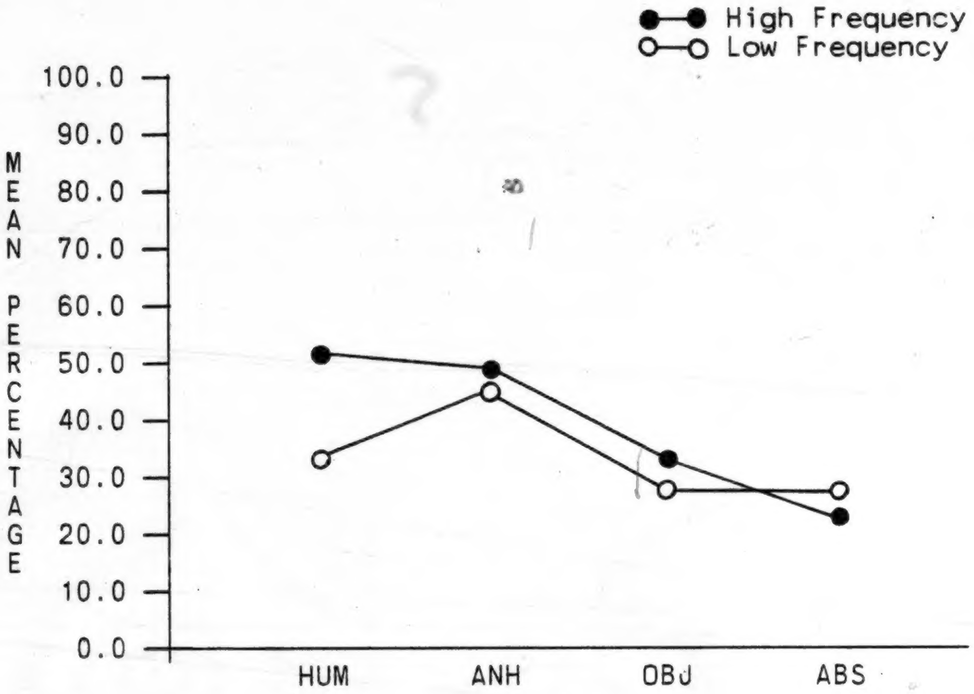


Figure 3. Interaction between Animacy-Concreteness and Frequency (HUM - Human, ANH - Animate Non-Human, OBJ - Object, and ABS - Abstract)

with regard to high frequency nouns, (i.e., the significant interaction between Frequency and Animacy - Concreteness).

The same series of Newman-Keuls *a posteriori* tests were performed with the percentage measurements as with the ratio measurements, so that the Frequency, Concreteness, Animacy Hypotheses could be investigated in more detail. That is, 28 possible mean differences were tested between the following eight cases: the Human, Animate Non-Human, Object, and Abstract cases of the High and Low Frequency conditions. The basic results are shown in Tables 7 and 8. (The detailed Q values and their significances are presented in Appendix E.)

Concerning the Frequency Hypothesis, the same results were obtained as in the case of the ratio measurements (see Table 7). That is, only the mean difference between the Human cases of the High and Low Frequency conditions was found to be significant ($Q5(87) = 7.087$, $p < 0.01$). The mean percentages for the Animate Non-Human and Object cases of the High Frequency condition turned out to be slightly higher than those for the Animate Non-Human and Object cases of the Low Frequency conditions, but the mean differences between these groups did not reach a significant level (see Tables 5 and 7). It follows, therefore, that the Frequency Hypothesis may be valid for human nouns, but not for other nouns.

Table 7

Results of *A Posteriori* Tests (Percentages)

Pairs	Sig.
HHUM - LHUM	***
HANH - LANH	
HOBJ - LOBJ	
HABS - LABS	

***$p<0.01$

Table 8

Results of *A Posteriori* Tests (Percentages)

	HHUM	HANH	HOBJ	HABS		LHUM	LANH	LOBJ	LABS
HHUM	—		***	***	LHUM	—	***		
HANH		—	***	***	LANH		—	***	***
HOBJ			—	*	LOBJ			—	
HABS				—	LABS				—

*$p<0.10$
***$p<0.01$

As the independent test of the Concreteness Hypothesis, the mean differences between the Object and Abstract conditions were tested. As a result, the mean difference between the Object and Abstract cases of the High Frequency condition, but not the Low Frequency condition, was found to be significant ($Q5(87) = 3.894$, $p < 0.10$). It is important to note that this mean difference was not shown to be significant with the use of the ratio measurements (see Tables 4 and 8). This result means that the independent validity of the Concreteness Hypothesis was supported using the percentage measurements, but not using the ratio measurements.

As far as the Animacy Hypothesis was concerned, the same results were obtained as those obtained with the use of the ratio measurements (see Tables 4 and 8). That is, two out of three mean differences, relevant to the independent test of the Animacy Hypothesis, were found to be significant in the case of the High Frequency condition: the Human and Object conditions, $Q4(87) = 6.594$, $p < 0.01$ and the Animate Non-Human and Object conditions, $Q3(87) = 5.997$, $p < 0.01$). It was also found that the mean percentage for the Human condition (i.e., 51.10) was higher than that for the Animate Non-Human condition (i.e., 49.46), but the mean difference between these two groups did not reach a significant level (see Tables 4 and 8). It can be argued, therefore, that the difference between animate and inanimate nouns is much bigger than the difference between human and animate non-human nouns with respect to the effects on the selection of grammatical subjects. This point is clearly consistent with what has been found with the use of the ratio measurements.

Furthermore, the interaction effects between the animacy and concreteness of high frequency nouns on the choice of grammatical subjects were found to be additive as in the case of the ratio measurements. This is because the mean difference between the Human and Abstract conditions turned out to be highly significant ($Q8(87) = 10.488$, $p < 0.01$).

Finally, the same results were found for the Low Frequency condition as those obtained on the basis of the ratio measurements (see Tables 4 and 8). More specifically, it turned out that low frequency animate non-human nouns were significantly more likely to be grammatical subjects than low frequency human, object, or abstract ones. This significant tendency is, of course, contrary to the expectation of the Animacy Hypothesis.

3.1.4 Discussion and Conclusions

It must be stressed first that two independent measurements produced very similar overall results. It is plausible, therefore, to regard these results as highly conclusive.

The Frequency Hypothesis was verified with regard to human nouns only. That is, high frequency human nouns were much more likely to be chosen as grammatical subjects in sentences than were low frequency human nouns. It was also found that high frequency object nouns were somewhat more likely to be grammatical subjects than low frequency object nouns, but the difference between these two groups did not reach a significant level. Contrary to the Frequency Hypothesis, low frequency animate non-human and abstract nouns were found to be somewhat more likely to be grammatical subjects than high frequency animate non-human and abstract nouns, respectively, but the mean differences between these two pairs of groups were not statistically significant (see Figures 1 and 3).

Since the main effect of Animacy - Concreteness turned out to be significant, it can be concluded that the Animacy Hypothesis was confirmed in conjunction with the Concreteness Hypothesis and vice versa. The point is, however, that this main effect holds true only for high frequency nouns, since the interaction effect between Frequency and Animacy - Concreteness was shown to be significant (see Figures 1 and 3).

As far as high frequency nouns were concerned, it was found that both the mean ratio and percentage for the Object condition were higher than those for the Abstract condition (see Figures 1 and 3). Furthermore, the mean percentage difference between these two groups reached a significant level, while the mean ratio difference did not (see Tables 4 and 8). Taken together, these results suggest that the independent effect of concreteness on the choice of grammatical subjects should hold only for high frequency nouns, at least as far as the percentage measurements are concerned.

On the other hand, it is rather surprising to find that low frequency abstract nouns turned out to be more likely to be grammatical subjects than low frequency object nouns (see Figures 1 and 3). This tendency is, of course, contrary to the Concreteness Hypothesis, but the mean ratio and percentage differences between these two groups did not reach a significant level.

The Animacy Hypothesis was clearly confirmed both dependently and independently of the Concreteness Hypothesis. It must be noted, however, that the Animacy Hypothesis holds only for high frequency nouns. This means that as far as high frequency nouns are concerned, the higher nouns are in the animacy hierarchy, the more likely they are to be chosen as the grammatical subjects in the sentences. As the results of the independent test of the Animacy Hypothesis, it can also be concluded that with regard to the choice of grammatical subjects, there is a significant cluster between human and animate non-human nouns on the one hand and object and abstract nouns on the other, with the former being more likely to be grammatical subjects than the latter (see Figures 1 and 3). This conclusion can be reached only for high frequency nouns.

On the basis of what has been found about the Concreteness and Animacy Hypotheses, the following hierarchy of the choice of grammatical subjects is proposed: (human > animate non-human) > object > abstract, where "A>B" means that A is more likely to be a grammatical subject than is B, (hierarchical differences inside the parentheses may not lead to statistically significant differences). Three points need to be made about this hierarchy. First, the first two types of nouns, human and animate non-human, are *much* more likely to be grammatical subjects than the other two types of nouns, object and abstract, regardless of whether ratio or percentage measurements are used. Second, object nouns are more likely to be grammatical subjects than abstract nouns, at least as far as percentage measurements are concerned. Third, this hierarchy must be regarded as being applicable only for high frequency nouns.

Similarly, the following hierarchy may be suggested for low frequency nouns: animate non-human > (human > abstract > object). It is important to notice that the overall results for low frequency nouns turned out to be totally contrary to all of the Frequency, Concreteness, and Animacy Hypotheses.

As an inherent property of nominal referents, animacy was found to be the most important non-discourse determinant of the choice of grammatical subjects in sentences. Frequency was also found to be an important non-discourse determinant in the sense that it interacted with other non-discourse determinants in interesting ways. For example, animacy and concreteness could be non-discourse determinants only with regard to high frequency nominal referents; furthermore, frequency is a major determinant with regard to human nouns, and plays a minor role in the case of object nouns.

Finally, some comments need to be made about the significant main effect of Task. The important point is that the interactions between Task and either Frequency or Animacy - Concreteness did not turn out to be significant. Therefore, the main effect can basically be interpreted in a literal sense. In general, subjects were more likely to use personal pronouns, proper nouns, and so on rather than given nouns as grammatical subjects in clauses, when they were asked to compose passages than when they were asked to compose single sentences. Consequently, it is reasonable to assume that the ratio and percentage scores for the passage task tend to be lower than those for the single sentence task. The significant main effect of Task can then be attributed to such a tendency. The task difference did not have any effect on the overall relationship between the four non-discourse properties of nouns and their likelihood of being grammatical subjects.

3.2 *Experiment 2*

3.2.1 *Purpose*

The purpose of Experiment 2 was to test the Prototypicality Hypothesis. It predicts that the more prototypical a nominal referent is, the more likely it is to be chosen by a speaker as his cognitive point of departure and thus as the grammatical subject in the sentence.

3.2.2 *Method*

Subjects. Twenty three University of Alberta students, 20 females and three males, volunteered to participate in Experiment 2. All the subjects were taking an Introductory Linguistics course. They were all native speakers of North American English.

Materials. The task of Experiment 2 was simply for the subjects to indicate which of a pair of sentences they felt to be the better description of a simple picture in which two geometric figures were depicted. These two figures were designed to differ in their prototypicality as instances of geometric shapes. For this purpose, the following eight geometric shapes were selected as equally familiar to average university students: *triangle*, *square*, *circle*, *rectangle*, *pentagon*, *oval*, *star*, and *diamond* shapes.

Two versions of each of these eight shapes were designed, with one as a "prototypical" shape and the other as a "distorted" shape. According to Rosch (1973) and the results of pilot studies, the following distortion rule was imposed: half of the shape was replaced by either "curved" or "wavy" lines. The halves of the square, triangle, diamond, and star shapes were then replaced by curved lines; those of the rectangle, pentagon, oval, and circle shapes were replaced by wavy lines.

Twenty-three booklets were made, each of which had one instruction page and eight other pages. The size of each page was 8in. X 11in. At the top of each of the 10 pages, a pair of the prototypical and distorted geometric shapes was depicted either vertically or horizontally, with the space between them being 1.0 inch. These pairs of shapes were randomly determined. As a result, the following eight pairs of prototypical and distorted shapes were prepared: (rectangle, pentagon); (pentagon, rectangle), (square, triangle), (triangle, square), (diamond, oval), (oval, diamond), (circle, star), and (star, circle), where the first and second members indicate prototypical and distorted shapes, respectively. The order of the pages was randomized.

Furthermore, the last four pairs of figures were designed so as to differ in size as well as shape. The first members of the fifth and the six and the second members of the seventh and the eighth were smaller that the other members; the differences between the sizes of the two figures of each pair were approximately equalized among the four pairs. The reason for this manipulation will become clear below.

Each of the first four pairs of figures was depicted vertically at the top center of the page, with the spaces being 1.0 inch. Each of the next four pairs of figures was depicted horizontally in the same way. Finally, the distorted members of the first and second pairs were located beneath the prototypical ones; the distorted members of the third and fourth pairs were located above the prototypical ones; and the distorted and prototypical members of the fifth through the eighth appeared alternately in the left and right sides of the pages.

For the sake of convenience, these eight pairs of pictures are hereafter referred to as "UV1," "UV2," "MV1," "MV2," "US1," "US2," "MS1," and "MS2," where "V" and "S" indicate "Vertical" and "Size," and "U" and "M" indicate "Unmarked" and "Marked." The actual pictures are shown in Appendix E.

About 2 inches below each of these pictures, there was a pair of incomplete sentences with two blanks. There were two types of sentence pairs. For UV1, UV2, MV1, and MV2, the following pair was designed: *(a) There is* ______ *that is above* _____ and *(b) There is* _____ *that is below* _____ . For

US1, US2, MS1, and MS2, the following pair was designed: *(a) There is _____ that is bigger than _____* and *(b) There is _____ that is smaller than _____*, where "__" indicates a blank representing a missing noun phrase, which the subjects were asked to fill in in order to complete one of each pair. In the experiment, 2.5 inch-long underlines were used. These two types of incomplete sentence pairs were typed about 2 inches below the pictures, with the space between them being 1.5 inches.

It is important to notice that the pairs of *above* and *below* and *bigger* and *smaller* have so-called "unmarked" and "marked" members, respectively, according to Clark (1973) and Clark and Clark (1977). *Above* and *bigger* are generally conceived of as the unmarked members, while *below* and *smaller* are the marked ones. Moreover, it is generally assumed that when given the choice between unmarked and marked predicates to describe the state of affairs, people tend to choose the unmarked ones rather than the marked ones. If this is so, the sentences with "above" and "bigger" might be chosen by the subjects rather than those with "below" and "smaller," respectively. That is, the subjects may choose the unmarked sentences rather the marked sentences.

If the Prototypicality Hypothesis is correct, the subjects may be expected to choose the marked sentences in the cases of MV1, MV2, MS1, and MS2. This is because in the

MV1 and MV2, the prototypical shapes (i.e., triangle and square ones) were located below the distorted ones, and in the MS1 and MS2, the prototypical shapes (i.e., star and circle ones) were smaller in size than the distorted ones. In other words, if the subjects choose to realize the prototypical shapes rather than the distorted ones as grammatical subjects in the sentences, the sentences will necessarily have the marked predicates, *below* and *smaller*.

It is important to raise the question of what the grammatical subject is in an English existential sentence such as *There is a circle that is above a triangle* (Quirk & Greenbaum, 1973). That is, which of "there" or "a circle" is a grammatical subject or a notional (or logical) subject? As far as the present experiments are concerned, grammatical subjects are defined operationally (and tentatively) as referring to nominal referents, for example "a circle" rather than "there" in the above sentence.

Procedure. The experiment was carried out in a single classroom, after the subjects had finished a class. Each subject was provided with one booklet, and then asked to read the instructions on the first page. They were then asked to proceed with the task if they were clear about the instructions. The task was for the subjects to choose one from each pair of incomplete sentences described above which they felt to be the better sentence for describing the picture, and then to complete it by filling in the two missing noun phrases.

In the instructions, the following points were stressed. First, the eight pictures were independent of each other, so that the subjects were supposed to ignore all others, while they were focusing on one particular picture. Second, there were no correct and incorrect sentences, and therefore the subjects were encouraged to choose one from each pair which they felt to be the better. Third, they were advised to make guesses if they had difficulty in identifying and naming some shapes. On the average, it took the subjects about 15 minutes to complete the task.

3.2.3 Analyses and Results

Some data were first excluded from further analysis for the following reasons. First, two subjects failed to understand the instructions; they did not choose one from each pair of sentences, completing both members of the pair. There were six cases in which subjects identified prototypical shapes incorrectly, for example the prototypical diamond shape was referred to as "a square." Furthermore, there were four "logically" incorrect cases in which subjects correctly named the shapes, but made an incorrect choice between "above" and "below" or between "bigger" and "smaller." For example they stated "There is a triangle that is above a square" for the picture in which the square was above the triangle.

Generally, as expected, it was found that the subjects had no difficulty in identifying and then naming the prototypical shapes, but they had a certain amount of difficulty with the distorted ones. The phrases referring to the prototypical shapes were found to be much shorter and to have much less variation than those referring to the distorted ones. Whatever phrases the subjects used in order to refer to the distorted shapes, they were counted as the noun phrases for the shapes, as long as they could be viewed as different from those of the prototypical shapes.

For the sake of numerical analysis, the choice of a marked sentence was assigned "one" and that of an unmarked one was assigned "zero." Either one or zero was then assigned to all the choices of the sentences the subjects made. These quantities will hereafter be referred to as "markedness scores." For each of the eight pictures, the markedness scores were obtained. High markedness scores indicate that the subjects tended to choose the marked sentence members, which have either the predicate "below" or "smaller." On the other hand, low markedness scores indicate that the subjects tended to choose the unmarked sentence members with either the predicate "above" or "bigger."

The analysis was focused on the following four factors which may have to do with the choice of grammatical subjects and then the choice between "above" and "below" and between "bigger" and "smaller." The four were the prototypicality of shapes (i.e., the Prototypicality Hypothesis), a "top-to-bottom" order principle, a "left-to-right" order principle, and the markedness of the predicates. If the prototypicality of shapes is a major factor for the choice of grammatical subjects and then of the choice between the two pairs of predicates, i.e., if the Prototypicality Hypothesis is confirmed, scores of MV1, MV2, MS1, and MS2 might be higher than those of UV1, UV2, US1, and US2. Since in MV1 and MV2 the prototypical shapes were located *below* the distorted ones, and the subjects were expected to choose the former as grammatical subjects, the marked sentences with "below" were expected to be chosen rather than the marked sentences with "above." Similarly, in the MS1 and MS2, the prototypical shapes were designed to be *smaller* than the distorted ones; therefore, the marked sentence with the predicate "smaller" may be expected to be chosen by the subjects.

The top-to-bottom order principle predicts that the subjects will place upper figures, whether they are prototypical or distorted ones, in earlier surface positions in the sentences than lower ones. If this is so, given a picture in which a figure A is located above another figure B, they might produce the unmarked sentence *There is A that is above B* rather than the other marked sentence. This top-to-bottom order predicts the low markedness scores of

UV1, UV2, MV1, and MV2. According to the left-to-right order principle, the figures on the left sides of pages may be more likely to be placed in earlier surface sentence positions than the figures on the right sides of the pages. The subjects were then expected to produce *There is A that is bigger/smaller than B* for a picture in which figure A was to the left of figure B. This left-to-right order principle predicts high markedness scores of US2 and MS2 and low markedness scores of US1 and MS1, since in the former pictures, the smaller figures were located to the left of the bigger ones. Finally, according to the markedness of the predicates, low markedness scores were expected for all the eight pictures.

The predictions of these four factors and the mean markedness scores of the pictures are as shown in Table 9. In order to test the Prototypicality Hypothesis, either one- or two-tailed *Sign* tests were performed so as to determine whether the differences between each of the following pairs of markedness scores were at a chance level or not: six pairs among UV1, UV2, MV1, and MV2 and six pairs among US1, US2, MS1, and MS2. If it was possible to expect one of each pair to produce higher markedness scores than the other, the directional one-tailed tests were performed, for example with UV1-MV1; otherwise non-directional two tailed tests were carried out, for example with UV1-UV2, in both of which none of the three factors predicts the choice of "above" (see Table 9).

As expected, the following four out of the first six pairs turned out to be significant: UV1-MV1, $Z = 1.77$, $p < 0.05$, UV1-MV1, $Z = 2.00$, $p < 0.05$, UV2-MV1, $Z = 2.00$, $p < 0.05$, and UV2-MV2, $Z = 2.21$, $p < 0.05$. Moreover, the following four out of the second six pairs were found to be significant: US1-US2, $Z = 2.04$, $p < 0.05$, US1-MS2, $Z = 2.71$, $p < 0.01$, US2-MS1, $Z = 2.48$, $p < 0.01$, and MS1-MS2, $Z = 2.85$, $p < 0.01$. These results are shown in Table 10.

From these, it follows that the Prototypicality Hypothesis was clearly confirmed only with regard to the choice between "above" and "below." Note that the mean markedness scores of MV1 and MV2 are 0.444 and 0.529; on the other hand, those of UV1 and UV2 are 0.059 and 0.059. As shown above, the four pairs between UV1 and UV2 on the one hand and MV1 and MV2 on the other hand were found to be significant. As can be seen in Table 9, only the Prototypicality Hypothesis predicted the choice of "below" over "above" in the case of MV1 and MV2. This means that the high mean markedness scores of MV1 and MV2 can be attributed only to the force of the factor of the prototypicality of shapes (i.e., the Prototypicality Hypothesis).

The Prototypicality Hypothesis was not confirmed, however, with regard to the choice between "bigger" and "smaller." The markedness score of MS2 turned out to be high as expected; however, that of MS1 was even lower than

Table 9

Predictions of the Three Factors and Mean Markedness Scores and Standard Deviations (Experiment 2)

	Prot.	T-B[1]	Mark.	Mean	SD
UV1	above	above	above	0.059	0.243
UV2	above	above	above	0.059	0.243
MV1	below	above	above	0.444	0.511
MV2	below	above	above	0.529	0.514

	Prot.	L-R[2]	Mark.	Means	SD
US1	big	big	big	0.125	0.342
US2	big	small	big	0.643	0.497
MS1	small	big	big	0.111	0.323
MS2	small	small	big	0.706	0.470

[1] Top-Bottom ordering principle
[2] Left-Right ordering principle

Table 10

Results of *Sign* Tests on Six Pairs (Experiment 2)

	UV1	UV2	MS1	MS2		US1	US2	MS1	MS2
UV1	—		*	*	US1	—	*		**
UV2		—	*	*	US2		—	**	
MV1			—		MS1			—	**
MV2				—	MS2				—

*$p<0.05$
**$p<0.01$

that of US1, which was totally contrary to the Prototypicality Hypothesis (see Table 9). It was clear that the overall results should be explained in terms of the factor of the left-to-right order principle. This is because the mean markedness scores for US2 and MS2 were found to be significantly higher than those for US1 and MS1 according to the results of the above *Sign* tests; and the left-to-right ordering principle predicts the high markedness scores for both US1 and MS2 and the low markedness scores for both US2 and MS1 (see Tables 9 and 10).

3.2.4 Discussion and Conclusions

It was clearly shown that the Prototypicality Hypothesis held true for the choice between the unmarked and marked pair of "above" and "below." That is, when prototypical shapes were located below distorted ones, the subjects tended to choose the marked predicate "below." This is because they are likely to choose the prototypical figures as grammatical subjects in sentences. In this sense, it can be concluded that the Prototypicality Hypothesis was confirmd only in conjunction with the choice between "above" and "below."

However, concerning the choice between "bigger" and "smaller," the Prototypicality Hypothesis was not confirmed at all. This finding can be taken to mean that the factor of the prototypicality of shapes failed to be a major determinant of the choice of grammatical subjects in conjunction with the choice between "bigger" and "smaller."

Instead, it was found that as far as US1, US2, MS1, and MS2 were concerned, the subjects adopted a simple left-to-right order principle, placing the figures on the left sides of pages in the subject position. The choice between "bigger" and "smaller" was then made on the basis of the choice of the leftmost figures as grammatical subjects.

It may be that the left-to-right order principle was so dominant that the other factor, the prototypicality of shapes, were surpressed. This point can be seen in the low markedness score for MS1 in which the Prototypicality Hypothesis and the left-to-right order principle predict the high and low markedness scores, respectively. Actually, this tendency was found in pilot studies. Therefore, in order to test the Prototypicality Hypothesis, the force of the left-to-right order principle should be avoided. Following these considerations, the next experiment was designed and carried out.

3.3 Experiment 3

3.3.1 Purpose

The purpose of Experiment 3 was to test the Prototypicality Hypothesis while avoiding interference from the left-to-right order principle.

3.3.2 Method

Subjects. Fifteen University of Alberta students, 13 females and two males, volunteered to take part in Experiment 3. They all were taking an Introductory Linguistics Course. They all were native speakers of North American English.

Materials. The materials of Experiment 2 were used; the only difference was that the pairs of figures were depicted vertically, not horizontally. That is, the figures (i.e., geometric shapes) on the left sides of the original pictures were placed above those of the figures on the right sides of the original pictures.

Fifteen booklets were made which had one instruction page and four other pages. On the top center of each of the four pages, a pair of figures was depicted vertically, with the space between them being 1.0 inch; about 2.5 inches below them a pair of incomplete sentences was typed, with the space between them being 1.5 inches.

Procedures. Experiment 3 was carried out in a classroom, after the subjects had finished a class. The procedures were the same as those of Experiment 2.

3.3.3 Analyses and Results.

The following data were first excluded from further analysis. First, three subjects misunderstood the instructions; they completed both sentences of each pair. It is rather suprising that three out of the 15 subjects failed to understand the instruction to choose between a pair of sentences, but this does not appear to invalidate the conclusion of the present experiment. Second, there were four cases in which subjects could not identity the prototypical shapes correctly. There was one "logically" incorrect case in which subject named the shapes correctly, but made an incorrect choice of the predicate.

The same scoring method was carried out with these data. The predictions of the top-to-bottom order principle, the Prototypicality Hypothesis, and the markedness of predicates and the mean markedness scores for the four pictures are represented in Table 11.

With these data the same either one- or two-tailed *Sign* tests were carried out as with the data in Experiment 2. As shown in Table 12, the differences between each of the following two pairs turned out to be significant: US1-MS2, $Z = 2.00$, $p < 0.05$ and MS1-MS2, $Z = 2.04$, $p < 0.05$.

As expected, the mean markedness scores for US2 and MS1 turned out to be somewhat higher than that for US1, but neither of the differences between US1 and US2 and US1 and MS1 reached a significant level (see Tables 11 and 12). This result implies that neither of the two factors, the prototypicality of shapes and the top-to-bottom order principle, should be considered to be major determinants of the choice of grammatical subjects in conjunction with the choice between "bigger" and "smaller."

It was also found that there was an additive interaction between the two factors in respect to the choice of grammatical subjects and then to the choice between "bigger" and "smaller." This is because the mean markedness (i.e., 0.818) for MS2 turned out to be approximately the mean markedness (i.e., 0.375) for US2 plus the mean markedness (i.e., 0.333) for MS1; the relatively high markedness scors for US2 and MS1 should be attributed to the forces of the top-to-bottom order principle and the prototypicality of shapes, respectively; and the high markedness score for MS2 may be attributed to both the top-to-bottom order principle and the prototypicality of

Table 11

Predictions of the Three Factors and Mean Markedness Scores and Standard Deviations (Experiment 3)

	Prot.	T-B[1]	Mark.	Means	SD
US1	big	big	big	0.167	0.389
US2	big	small	big	0.375	0.518
MS1	small	big	big	0.333	0.492
MS2	small	small	big	0.818	0.405

[1] Top-Bottom ordering principle

Table 12

Results of *Sign* Tests on Six Pairs (Experiment 3)

	US1	US2	MS1	MS2
US1	—			*
US2		—		
MS1			—	*
MS2				—

$*p < 0.05$

shapes foctors (see Table 11).

In order to compare the strengths of the three factors, the prototypicality of shapes (i.e., the Prototypicality Hypothesis), the top-to-bottom order principle, and the left-to-right order principle, simple estimations were carried out using mean markedness scores obtained in both Experiments 2 and 3. The strength of the prototypicality of shapes was estimated by the mean of the following three estimates: ((MV1 - UV1) + (MV2 - UV2)) / 2, ((MS2 - US2) + (MS1 - US1)) / 2, and ((MS2 - US2) + (MS1 - US1)) / 2, where the mean markedness scores in the first and second are from Experiment 2; the mean markedness scores in the third from Experiment 3. It is important to notice that the result of each subtraction can be attributed only to the force of the prototypicality of shape, for example in the case of (MV1 - UV1), only the Prototypicality Hypothesis predicts the choice of "below" for MV1, while none of the three factors predicts the choice of "below" for UV1 (see Tables 9 and 11). Similarly, the strengths of the left-to-right and top-to-bottom order principles were estimated as follows: ((MS2 - MS1) + (US2 - US1)) / 2, using the scores from Experiment 2, and ((MS2 - MS1) + (US2 - US1)) / 2, using the scores from Experiment 3.

As a result, the following estimated values were calculated: 0.253 for the prototypicality of shapes, 0.347 for the top-to-bottom order principle, and 0.557 for the left-to-right order principle. From this, the following hierarchy can be suggested: the left-to-right order principle > the top-to-bottom order principle > the prototypicality of shapes, where "A>B" means that A is more dominant with respect to the choice of grammatical subejcts and then the choice between the pairs of unmarked and marked predicates.

3.3.4. Discussion and Conclusions

In Experiment 3, it was observed that the subjects tended to choose prototypical shapes as grammatical subjects rather than others, but this tendency failed to reach a statistically significant level. On the basis of the results obtained in Experiments 2 and 3, it can be concluded that the Prototypicality Hypothesis cannot be fully supported in conjunction with the choice between "bigger" and "smaller." This conclusion is puzzling in the light of the fact that the hypothesis was clearly verified in conjunction with the choice between "above" and "below."

It was also suggested that there is an additive interaction among the three factors, the prototypicality of shapes, the top-to-bottom order principle and the left-to-right order principle: that is, the three factors interact with each other in an additive manner with regard to the choice of grammatical subjects and the choice between pairs of unmarked and marked predicates. A simple example of the additive interaction between the prototypicality of shapes and the top-to-bottom order principle is this: when a prototypical smaller figure, A, is located above a distorted bigger figure, B, people may be more likely to choose the prototypical figure as a grammatical subject, (i.e., *There is A that is smaller than B)*, than when the prototypical smaller figure is located below the bigger distorted one.

Finally, it seems reasonable to suggest the following hierarchy of the three factors on the basis of their estimated strengths: the left-to-right order principle > the top-to-bottom order principle > the prototypicality of shapes, with regard to the choice of grammatical subjects and then to the choice between pairs of unmarked and markedness predicates. It should be noted that the left-to-right order principle can be viewed as a much stronger determinant of the choice of grammatical subjects and the choice between marked and unmarked pairs of predicates than either the prototypicality shapes or the top-to-bottom order principle. It is also true that these factors work in an additive manner as major or minor determinants of the choice of grammatical subjects.

4. GENERAL DISCUSSION AND CONCLUSIONS

In this chapter, the results obtained in Experiments 1, 2 and 3 will be discussed in comparison with several related psycholinguistic studies on the choice of grammatical subjects. Subsequently, some psycholinguistic implications will be offered for the notion of grammatical subject and for a nearly universal relative order of subjects before other nominal constituents. Finally, several limitations of this work will be discussed.

4.1 Summary and Discussion of Results

This work focuses on four inherent non-discourse properties of nominal referents as the major determinants of a speaker's cognitive point of departure which may be assumed to be a primary psychological function of a grammatical subject. The Frequency, Concreteness, Animacy, and Prototypicality Hypotheses have been empirically tested.

The Frequency Hypothesis was found to be true only for human nouns. That is, the more frequent a human noun is, the more likely it is to be chosen by a speaker as a grammatical subject. It also turned out that the subjects tended to realize high frequency object nouns as grammatical subjects rather than low frequency object nouns, regardless of whether they produced single sentences or passages. However, the differences between these high and low frequency object noun groups failed to reach a statistically significant level.

The results for animate non-human and abstract nouns failed to support the Frequency Hypothesis. That is, there was a tendency for low frequency animate non-human and abstract nouns to be grammatical subjects more often than high frequency ones, but this tendency was not statistically significant (see Figure 1 and Table 3). It may be speculated that since the subjects had some difficulty in formulating the contents of sentences using low frequency animate non-human and abstract nouns, they simply wrote "definitional" sentences which started with such phrases as "A dodo is ....," "Affluence is ...." and "Rationalization is ......"

With regard to the percentage measurements, the same overall results were found for the Frequency Hypothesis (see Figures 1 and 3). One noticeable point is that low frequency animate non-human nouns were found to be slightly less likely to be grammatical subjects than high frequency ones. Consequently, the main effect of Frequency turned out to be significant (see Figure 3 and Table 6).

Taken together, these observations suggest that high frequency human referents should be much more available in the mind of a speaker than low frequency human ones, with the former being more likely to be the cognitive point of departure. Furthermore, this holds true, to some extent, for object referents. The following conclusion can then be drawn: the more frequent a human referent is, the more likely it is to be a grammatical subject, and the frequency of other kinds of nominal referents does not cause significant effects on their likelihood of being grammatical subjects.

It is possible to attempt to account for the fact that only high and low frequency human nouns were treated as predicted by the Frequency Hypothesis. One possible explanation is that some of the low frequency human nouns, such as *buffoons*, *idlers* and *anarchists*, do not only refer to human referents, but also predicate some properties, qualities, or attributes to them. In other words, these nouns tend to function as nominal predicates, carrying more information about the predicative aspects of humans than high frequency human nouns, such as *boys*, *teachers* and *players*.[3] This point needs to be further investigated.

It appears that the frequency of nominal referents has been discussed in neither the linguistic nor psycholinguistic literatures as a possible determinant of subject selection in a sentence. The present work suggests that the frequency of nominal referents should be viewed as a major determinant of the selection of grammatical subjects, at least as far as human referents are concerned.

---

[3] I am indebted to both Dr. M. Dryer and Dr. G.D. Prideaux for this explanation.

The Concreteness Hypothesis was confirmed in conjunction with the Animacy Hypothesis and vice versa. That is, concrete nouns, whether human, animate non-human or object, were found to be more likely to be grammatical subjects than abstract ones. It is important, however, to note that this tendency holds only for high frequency nouns.

The independent test of the Concreteness Hypothesis focused on the differences between object and abstract noun groups. It turned out that as far as high frequency object and abstract nouns were concerned, the mean percentage difference was found to be significant, while the mean ratio difference was not (see Tables 4 and 8). It should be mentioned, however, that both the mean ratio and percentage for the object noun group tended to be higher than those for the abstract noun group, regardless of whether the subjects were asked to compose single sentences or passages (see Figures 1 and 3). Therefore, it can be claimed that high frequency object nouns are more likely to be grammatical subjects than high frequency abstract nouns, but the statistical significance of this tendency may be subject to the measurements used.

These findings suggest that high frequency concrete object referents may be somewhat more available in the mind of a speaker than high frequency abstract ones, with the former being more likely candidates as the cognitive points of departure. Therefore, it can be concluded that high frequency concrete nominal referents are more likely to be chosen by a speaker as grammatical subjects than high frequency abstract nominal ones. The concreteness of nominal referents can then be viewed as a major independent determinant of subject selection in a sentence, at least as far as high frequency nominal referents are concerned.

The overall results for low frequency objects and abstract nouns seemed to be inconsistent with the Concreteness Hypothesis. That is, low frequency abstract nouns were found to be slightly more likely to be grammatical subjects than low frequency object nouns, but the difference between these two noun groups did not reach a significance level (see Figures 1 and 3).

This clear lack of evidence for low frequency object and abstract nouns is rather surprising, given a number of psychological studies (e.g., James, 1972; Paivio, 1971; Paivio & Begg, 1981) on the imagery-concreteness of nouns (see Section 2.2). In particular, James, Thompson, and Baldwin (1973) reported that when their subjects were asked to recall English active sentences in which the grammatical subjects and objects had low and high imagery-values, respectively, they changed the voice at recall 24% of the time. On the other hand, when the subjects were supposed to recall active sentences whose grammatical subjects and objects had high and low imagery-values, the change of voice took place at recall only 2.9% of the time. James, Thompson, and Baldwin then concluded that one major function of English passive sentences was to place high imagery-value nouns before low imagery-value ones.

Clearly, this conclusion is questionable, given the findings of Experiment 1, which suggest that the effects of the imagery-concreteness of nouns on subject selection should be limited to high frequency nouns. It is important to note that James, Thompson, and Baldwin did not take into account the frequency and animacy of nouns as possible psychological factors which might be related to the phenomena. That is, what they called "effects of imagery on the choice of voice" could possibly be related to three noun properties: concreteness (object vs. abstract), animacy (human, animate non-human, vs. inanimate), and frequency (high vs. low). It seems reasonable to suggest that if they had taken into account the frequency of nouns, their overall results might have come out differently, for example as suggested in this work. Furthermore, James, Thompson and Baldwin clearly failed to realize the independent aspects of two inherent properties, concreteness and animacy; for example, the referents of human nouns are not only concrete, but also animate, while the referents of object nouns are concrete, but not animate.

The Animacy Hypothesis was confirmed in a clear-cut manner, but interestingly enough, it held true only for high frequency nouns. This was confirmed using both ratio and percentage measurements (see Figures 1 and 3). That is, the higher the high frequency nouns were in the animacy hierarchy, the more likely they were to be grammatical subjects.

In particular, the *a posteriori* tests showed that there was a highly significant clustering between human and animate non-human nouns on the one hand and object and abstract ones on the other, with the former being much more likely to be grammatical subjects than the latter (see Tables 4 and 8). However, the differences between human and animate non-human noun groups did not turn out to be significant, although the mean ratios and percentages for human nouns were always higher than those for animate non-human ones (see Tables 1 and 5). These findings indicate that the distinction between animate and inanimate referents should be considered to be a much more important factor in subject selection than the distinction between human and animate non-human ones.

Concerning low frequency nouns, the overall results of Experiment 1 did not seem to be consistent with the Animacy Hypothesis. More specifically, low frequency human and object nouns were shown to be less likely to be grammatical subjects than low frequency animate non-human and abstract nouns, respectively (see Figures 1 and 3).

Given these findings, it may be argued that as far as high frequency nouns are concerned, the higher nominal referents are in the animacy hierarchy, the more available they are in the minds of speakers and the more likely they are to be chosen as their cognitive points of departure. Therefore, it can be concluded that the higher a high frequency nominal referent is in the animacy hierarchy, the more likely it is to be chosen by a speaker as a grammatical subject. Animacy as an inherent non-discourse property of nominal referents can be regarded as an important determinant of the selection of grammatical subjects.

This conclusion should hardly be surprising, given that a number of both linguistic and psycholinguistic studies (e.g., Dewart, 1979; Givón, 1976; Tomlin, 1979) argue for a strong association among grammatical subjects, animacy of nouns, and the semantic role of agent. It has been argued that grammatical subjects tend to refer to the semantic role of agent rather than other semantic roles, and the notion of agentivity is positively correlated with the animacy of nouns, especially with the distinction between animate and inanimate referents (see Section 2.2.1). The point is that this work provided empirical evidence in support of a formal linguistic argument. More specifically, the tendency for the semantic roles of agent, experiencer, and/or beneficiary to be grammatical subjects may be explained in terms of their conceptual properties of animacy and/or concreteness (Itagaki & Prideaux, 1984).

It must be made clear, however, that as far as low frequency nouns are concerned, animacy can not be viewed as a major non-discourse determinant of the selection of grammatical subjects. This means that the association among grammatical subjects, the animacy of nouns, and the semantic role of agent must be limited to high frequency nouns or nominal referents. It appears that animacy as a major non-discourse determinant of grammatical subjects has never been discussed in the linguistic and psycholinguistic literatures in connection with its interaction with the determinant, namely frequency.

On the grounds of the findings of Experiment 1, the following hierarchies of nouns with regard to subject selection can be concluded: for low frequency nouns, animate non-human > (human > abstract > object); and for high frequency nouns, (human > animate non-human) > object > abstract, where "A>B" means that A is more likely to be a grammatical subject than is B, (hierarchical relations inside the parentheses may not be considered to be statistically significant differences.)

Two points must be made clear about the second hierarchy. First, animate referents, whether humans or non-humans, are much more likely to be grammatical subjects than inanimate referents, regardless of whether people produce single sentences or passages. This tendency is also true, whether the ratio or percentage measurements are used. Second, there is no doubt about the tendency for object nouns to be grammatical subjects more often than abstract ones, but the hierarchical difference between these two groups may or may not be statistically significant.

It is generally maintained that as mentioned earlier, the tendency for animate referents to be grammatical subjects may be attributed to the strong association between the agentivity and animacy of the referents. There is another possible explanation. That is, it is possible to argue that as one universal aspect of human languages, people tend to develop and acquire more lexical verbs associated with animate entities than inanimate ones, and verbs which co-occur with animate agents/subjects more than with non-agents (Clark & Begun, 1971). This tendency may have been reflected, to some extent, in the results.[4] It is clear, however, that these two explanations cannot account for the finding that low frequency human and abstract nouns turned out to less likely to be grammatical subjects than low frequency animate non-human and object nouns, respectively. That is, low frequency nouns are not consistent with the animacy-concreteness hierarchy of nouns and the lexical structure with regard to subject selection. This observation differentiates the hypotheses of the present work from those of Clark and Begun, with the latter refuted by the present results.

---

[4] Initially, Dr. M. Dryer brought my attention to this point.

Finally, in Experiment 2, the Prototypicality Hypothesis was confirmed. That is, the subjects tended to choose more prototypical (or identifiable) geometric figures as grammatical subjects than less prototypical ones. It was also shown that this tendency was not the case when the subjects could adopt a simple "left-to-right" order principle. This means that when the subjects were asked to describe simple pictures in which two figures were depicted horizontally, they simply placed the figures on the left sides of the pictures in earlier surface positons than those on the right sides. (This tendency was found in pilot studies.) Furthermore, the force of the left-to-right order principle was found to be stronger than that of the prototypicality of shapes, with the latter surpressed by the former. On the basis of the results of Experiments 2 and 3, it can be concluded that the prototypicality of shapes may play a major role in the choice of grammatical subjects, but that role is not as dominant as that played by the left-to-right order principle (Tables 9-11).

Furthermore, it can be concluded that more prototypical (or identifiable) nominal referents are more available in the mind of a speaker than less prototypical nominal ones, and therefore more likely to serve as the cognitive point of departure. From this, the following conclusion can be drawn: the more prototypical a nominal referent is, the more likely it is to be chosen by the speaker as a grammatical subject. The prototypicality of nominal referents can in turn be viewed as a major non-discourse determinant of the choice of grammatical subjects.

Interestingly enough, it was found that the three factors, the prototypicality of shapes, the left-to-right order principle and the top-to-bottom order principle, interacted with each other additively with respect to subject selection. Moreover, the following hierarchy of the strengths of the three factors may be suggested, based on the mean markedness scores obtained in Experements 2 and 3: the left-to-right order principle > the top-to-bottom order principle > the prototypicality of shapes, where "A>B" indicates that A plays a more dominant role in subject selection than does B.

These findings seem to be somewhat contrary to what Johnson (1968a, 1968b, 1977) and Levelt and Massan (1981) found. First, these authors reported that the left-to-right order principle was not found to be the force which governed the choice of word order in sentences to be produced. For example, Levelt and Massan asked subjects to describe a simple film in which two geometic figures (for example, one triangle and one circle) were moving up together by choosing between a pair of the following sentences: *A triangle and a circle go up* and *A circle and a triangle go up*. Notice that the orders of the subject noun phrases are reversed. Levelt and Massan found that the subjects did not adopt the left-to-right order principle in choosing one from the pair; that is, the subjects did not simply choose the sentence in which the figure on the left side of the picture was placed in the first position in the conjoined noun phrase. Johnson (1968a, 1968b) also reported that he did not find any evidence for the left-to-right order principle as such (see Section 2.2.2).

Second, Levelt and Massan (1981) argue that the naming difficulty of nominal referents does not cause major effects on word order in sentences. They varied the naming difficulty of some, not all, geometric figures by forcing their subjects to refer to them by different nouns, for example they asked the subjects to refer to a triangle figure as either "tent" or "roof." Levelt and Massan assumed that if the subjects were forced to refer to such simple geometric figures by "inappropriate" nouns, they might have a certain amount of difficulty in naming them. They then hypothesized that subjects might place nouns with less naming difficulty in earlier surface positions in sentences than those with more naming difficulty. Their hypothesis was not confirmed. It was instead found that the subjects selected figures with more naming difficulty as the first noun phrase of the conjoined subject phrase 70% of the time, contrary to their expectation.

Levelt and Massan's (1981) findings may lead us to argue that the naming difficulty, codability, or prototypicality of nominal referents do not cause any effects on word order phenomena (cf. Bock, 1982; Bock & Irwin, 1980). One critical point must, however, be made about Levelt and Massan's conclusion. Their operational manipulation of naming difficulty may not be considered to be such that people may have more difficulty in naming some nominal referents than others. If people are forced to refer to a "triangle" figure as "tent" or "roof," they may not have any difficulty with it, since tents and roofs can be easily symbolized as the geometric shapes of triangles. Furthermore, they may have some affective bias toward the tents or roofs, which is similar to Osgood and Bock's (1977) notion of "vividness salience" (also see Ertel, 1977). It can be argued that since the subjects might have some kind of affective bias towards tents and roofs and against simple geometric figures, they placed the "inappropriate nouns" in the first position of the conjoined noun phrase 70.5% of the time. In this sense, it does not seem reasonable to argue that the naming difficulty, codability, or prototypicality of nominal referents may not be viewed as one possible determinant of the selection of grammatical subjects.

The results of Experiments 2 and 3 can be related to MacWhinney's (1977) view of subject selection. He first identifies three factors which he assumes to play major roles in the choice of grammatical subjects: "perceptual," "relational" and "linguistic." In discussing the perceptual factor, MacWhinney assumes that nominal referents in a sentence differ in the degree of perceptual salience (e.g., size, brightness, etc.), and that nominal referents with high degree of perceptual salience are likely to be grammatical subjects. The relational factor had to do with a speaker's preference for unmarked predicates over marked ones, for example *above* and *front* over *below* and *behind*, respectively. The point is that speakers tend to choose unmarked members rather than marked ones, with grammatical subjects in turn determined by that choice. Finally, the linguistic factor refers to the contextual aspects of nominal referents, for example "given" information, "topic," and so on. It is likely for speakers to choose given information as a grammatical subject rather than new information.

The present work offers some psycholinguistic content for the perceptual factor. That is, it seems possible to relate the perceptual factor to the three psychological elements, the left-to-right order principle, the top-to-bottom order principle and the prototypicality of shapes.

4.2 Psycholinguistic Contributions to Linguistics

As stated earlier, the present work has two psycholinguistic implications for linguistics. One had to do with Keenan's (1976) work on a "universal definition of subject." The other is the possibility of offering a partial psycholingusitic account for a nearly universal relative order of subject before other nominal constituents. (This order will hereafter be referred to as "SO" order.)

The results of this work can be related to two characteristics, "independent existence" and "autonomous reference," which are among a number of characteristics of the subjects in basic sentences, namely "universal" sets of characteristics (see Section 2.3). These two characteristics are also related to the traditional logical distinction between "particular (or singular)" terms and "universal (or general)" terms (Lyons, 1968; Sandmann, 1979). The particular terms consist of proper names, pronouns, and so on which denote individual *substances*. On the other hand, the universal terms are composed of common and abstract nouns, adjectives, and adverbs, which denote the *accidental* properties (e.g., quantity, quality, relation, action, place, state, etc.) of the individual substances. Given one particular term, for example *Socrates*, and one universal term, for example *man*, a sentence subject should be the particular one, namely

Socrates is a man. Of two universal terms, for example *Men* and *wise*, the more specific term should be a sentence subject, namely *Men are wise*.

The point is that according to Lyons (1968), these characteristics have traditionally been considered to be non-discourse determinants of grammatical subjects. Moreover, these characteristics are clearly related to the conceptual animacy hierarchy and/or concreteness of nominal referents (see Section 2.3). In other words, animate and/or concrete entities may be considered to be more independent entities, to be more referential, and to be more substantial than inanimate and/or abstract ones.

In this work, it was found that the Animacy and Concreteness Hypotheses were confirmed only with regard to high frequency nouns. Therefore, the association between the animacy of nominal referents and grammatical subjects should be restricted to high frequency nominal referents. It can also be suggested that the frequency of nominal referents should be taken into account as a major characteristic of grammatical subjects, at least as far as the psychological aspects of the grammatical subjects are concerned. It should be noted that the frequency of nominal referents was not discussed as such in Keenan's (1976) universal characteristics.

Finally, it is possible to discuss a nearly universal SO order in connection with the findings made in this work.

It was found that the four inherent non-discourse properties of nominal referents could be related to the selection of grammatical subjects as either major or minor determinants. Frequency was found to be a major determinant with regard to human and object referents; animacy is also a major determinant in conjunction with concreteness and vice versa; but they are limited to high frequency nominal referents, and therefore may not be major determinants for low frequency nominal referents. Prototypicality is also a major determinant, but it is not as dominant as the left-to-right order principle.

From these, it can be concluded that grammatical subjects tend to refer to nominal referents of high availability in human cognition rather than those of low availability. This is the first part of the possible partial psycholinguistic account for the dominant SO order. The second part is that according to Bock (1982), the easier nominal referents are to lexicalize, the earlier they are produced during sentence production. There is no doubt that the high cognitive availability of nominal referents is among the possible psychological elements which contribute to their ease of lexicalization.

It is suggested here that grammatical subjects tend to refer to nominal referents of high cognitive availability rather than those of low cognitive availability, and that the high cognitive availability of the subjects contributes to ease of lexicalization, which in turn results in these nouns occuring earlier in surface positions than others. As discussed in section 2.3, there is another possible psychological account for the dominant SO order. That is, grammatical subjects tend to refer to given information rather than new information, and speakers tend to place given information in earlier surface positions than new information (e.g., Bock, 1977; Bock & Irwin, 1980; Clark & Clark, 1978). These two explantions could, therefore, be viewed as working in tandem to provide psychological motivation for the linguistic fact that subjects tend to occur in earlier sentence positions than other noun phrases. That is, sentence-initial positions of subjects may be due both to the tendency for given information to precede new information and to the tendency for subjects to refer to nominal referents with high lexicalizing ease.

A psycholinguistic account of the dominant SO order is not so trivial in the linguistic literature. A number of linguists (e.g., Comrie, 1981; Dik, 1978; Givón, 1979) agree that certain psychological factors or constraints of speakers and hearers must be manifested in the structures of languages, and that certain linguistic phenomena must be explained ultimately in terms of psychological elements. The dominant SO order is among these linguistic phenomena which have been discussed in terms of certain psychological elements, but has never been explored as such in the linguistic literature. In this respect, Comrie (1981) states:

> Explanations for the predominance of word orders where the subject precedes the object seem more likely to have *a psychological basis*, in terms of the salience of the agent in the agent - action - patient situation, and the high correlation between semantic agent and syntactic subjects ... (emphasis added, p. 20)

Furthermore, Comrie explains the notion of salience as follows:

> It is, however, possible that certain language universals can be correlated with other aspects of *human cognitive psychology*, that are amenable to independent testing. .... we shall see that a certain hierarchy of noun phrases, which has significant relevance to cross - language generalization, also correlates highly with *an independently verifiable hierarchy of salience of entities in perception*." (emphases added, p. 25)

The present work has clearly demonstrated the possibility of accounting for the dominant SO order in terms of psychological elements, specifically the inherent non-discourse properties of nominal referents. This is only an initial step in a direction suggested but as yet unexplored by a number of *descriptive* linguists, such as those mentioned above.

4.3 Some Limitations of the Present Thesis

Finally, it is necessary to mention the limitations of this work. It may be suggested that the cognitive availability of nominal referents should be defined in terms of certain psychologically measurable variables rather than in terms of inherent nominal properties. It is possible, for example, to define the cognitive availability of nouns using reaction time data from studies in which subjects must distinguish real words from non-words. The more time is necessary in such lexical decisions, the less cognitive access there would be (Forster, 1976; Forster & Chambers, 1973). It would be interesting to investigate the correlation between the cognitive availability of nominal referents, defined in such a way, and their surface positions in sentences (Bock, 1982).

The present work sought to contribute to a psycholinguistic account for a nearly universal relative order of grammatical subjects before other nominal constituents. However, the experiments in this work were focussed only on subject selection in English sentences. There is no question, therefore, that the hypotheses tested here should be replicated in languages other than English; paticularly languages having relatively free basic word orders. It would be particularly interesting to determine whether animacy and/or concreteness play a major or minor role in subject selection in other languages.

One striking fact to come out of the Experiment 1 was that low frequency human and object nouns were shown to be less likely to be grammatical subjects than low frequency animate non-human and abstract nouns. This finding is of great interest, given a well-established linguistic tendency for nouns high in an animacy hierarchy to be grammatical subjects (Comrie, 1981; Tomlin, 1979). It is important, therefore, to investigate further the noun frequency in terms of how it interacts with other major or minor determinants of the choice of grammatical subjects. As far as low frequency nouns are concerned, the experimental results in this work lead us to conclude that an animacy hierarchy does not play any role in subject selection in English sentences.

REFERENCES

Anderson, J.R. *Cognitive psychology and its implications.* San Francisco: W.H. Freeman and Company, 1980.

Anderson, J.R., and Bower, G.H. *Human associative memory.* Washington, D.C.: Winston, 1973.

Allerton, D.J. The notion of 'givenness' and its relations to presupposition and to theme. *Lingua*, 1978, *44*, 122-168.

Allerton, D.J. *Essentials of grammatical description.* London: Routledge and Kegan Paul, 1979.

Bartlett, F.C. *Remembering: A study of experimental and social psychology.* Cambridge: Cambridge University Press, 1932.

Battig, W.F., and Mantague, W.E. Category norms of verbal items in 56 categories: A replication and extension of Connecticut category norms. *Journal of Experimental Psychology, Monograph*, 1969, *80*, 1-46.

Berlin, B., and Kay, P. *Basic color terms: Their universality and evolution.* Berkeley: University of California Press, 1969.

Berlyne, D.E. *Conflict, arousal, and curiosity.* New York: McGraw-Hill, 1960.

Black, M. *Language and philosophy.* Ithaca: Cornell University Press, 1949.

Bock, J.K. the effect of a pragmatic presupposition on syntactic structure in question answering. *Journal of Verbal Learning and Verbal Behavior*, 1977, *16*, 723-734.

Bock, J.K., and Irwin, D.E. Syntactic effects of information availability in sentence production. *Journal of Verbal Learning and Verbal Behavior*, 1980, *19*, 467-484.

Bock, J.K. Toward a cognitive psychology of syntax: Information processing contribution to sentence formulation. *Psychological Review*, 1982, *89*, 1-47.

Bolinger, D.L. *Aspects of language* (3rd, ed.). New York: Harcourt Brace Jovanovich, 1981.

Brown, R.W. *Words and things*. Glencoe, Illinois: Free

Press, 1958.

Brown, R.W. Reference: In memorial tribute to Eric Lenneberg. *Cognition*, 1976, *4*, 125-153.

Brown, R.W., and Lenneberg, E.H. A study in language and cognition. *Journal of Abnormal and Social Psychology*, 1954, *49*, 454-462.

Caramazza, A., and Brones, I. Semantic classification by bilinguals. *Canadian Journal of Psychology*, 1980, *34*, 77-81.

Carroll, J.B., Davies, P., and Richman, B. *Word frequency book*. New York: Houghton Mifflin Company, 1971.

Carroll, J.B., and Whilte, M.N. Word frequency and age of acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology*, 1973, *25*, 88-95.

Chafe, W.L. *Meaning and the structure of language*. Chicago: University of Chicago Press, 1970.

Chafe, W.L. Givenness, contrastiveness, definiteness, subject, topic, and point of view. In C.N. Li (Ed.), *Subject and topic*. New York: Academic Press, 1976.

Chafe, W.L. Creativity in verbalization and its implications for the nature of stored knowledge. In R.O. Freedle (Ed.), *Advances in discourse processes* (Vol. 1). Norwood, New Jersey: Ablex, 1977.

Chomsky, N. *Aspects of the theory of syntax*. Cambridge, Massachusetts: The M.I.T. Press, 1965.

Clark, E.V., and Clark, H.H. Universals, relativity, and language processing. In J.H. Greenberg (Ed.), *Universals of human language* (Vol. 1). Stanford: Stanford University Press, 1978.

Clark, H.H. Space, time, semantics, and the child. In T.E. Moore (Ed.), *Cognitive development and the acquisition of language*. New York: Academic Press, 1973.

Clark, H.H., and Begun, J.S. The semantics of sentence subject. *Language and Speech*, 1971, *14*, 34-46.

Clark, H.H., and Chase, W.G. Perceptual coding strategies in the formation and verification of description. *Memory and Cognition*, 1974, *2*, 101-111.

Clark H.H., and Clark, E.V. *Psychology and language*. New

York: Harcourt Brace Jovanovich, 1977.

Clark, H.H., and Haviland, S.E. Comprehension and the given-new contract. In R.D. Freedle (Ed.), *Advances in discourse processes* (Vol. 1). Norwood, New Jersey: Ablex, 1977.

Comrie, B. *Language universals and linguistic typology*. Oxford: Basil Blackwell, 1981.

Costerman, J., and Hupet, M. The other side of Johnson-Laird's interpretation of the passive voice. *British Journal of Psychology*, 1977, *68*, 107-111.

Creider, C.A. On the explanation of transformation. In T. Givón (Ed.), *Syntax and semantics* (Vol. 12). New York: Academic Press, 1979.

Cruse, D.A. Some thoughts on agentivity. *Journal of Linguistics*, 1973, *9*, 11-23.

Dewart, M.H. Role of animate and inanimate nouns in determining sentence voice. *British Journal of Psychology*, 1979, *70*, 135-141.

Dik, S.C. *Functional grammar*. Amsterdam: North-Holland Publishing, 1978.

Dik, S.C. *Studies in functional grammar*. New York: Academic Press, 1980.

Dik, S.C. *Two constraints on relators and what they can do for us*. In S.C. Dik (Ed.), *Advances in functional grammar*. Dordrecht-Holland: Foris Publications, 1983.

Earlhard, B. Determinants of the word-frequency effect in recognition memory. *Memory and Cognition*, 1982, *10*, 115-124.

Engelkamp, J., and Zimmer, H.D. *Dynamic aspects of language processing: Focus and presupposition*. Berlin: Springer-Verlag, 1983.

Ertel, S. Where do the subjects of sentences come from? In S. Rosenberg (Ed.), *Sentence Production: Developments in research and theory*. Hillsdale, New Jersey: Lawrence Erlbaum Assocites, 1977.

Fillmore, C.J. An alternative to checklist of meaning. *Proceedings of the first annual meeting of the Berkeley Linguistic Society*. Berkeley: Berkeley Linguistic Society, 1975.

Fillmore, C.J. The case for case reopened. In P. Cole and

J.M. Sadock (Eds.), *Syntax and semantics* (Vol. 8). New York: Academic Press, 1977.

Firbas, J. On defining the theme in functional sentence analysis. *Travaux Linguistiques de Prague*, 1964, *1*, 267-280.

Firbas, J. Non-Thematic subjects in English. *Travaux Linguistiques de Prague*, 1966, *2*, 239-256.

Firbas, J. Some aspects of the Czechoslovak approach to problems of functional sentence perspective. In F. Danes (Ed.), *Papers on functional sentence perspective*. The Hague: Mouton, 1974.

Flavell, J.H. *The developmental psychology of Jean Piaget*. Princeton, New Jersey: Van Nostrand Reinhold, 1963.

Flores d'Arcais, G.B. Some perceptual determinants of sentence construction. In G.B. Flores d'Arcais (Ed.), *Studies in perception*. Milan, Italy: Martello-Guninti, 1975.

Forster, K.I. Accessing the mental lexicon. In E. Walker and R. Wales (Eds.), *New approaches to language mechanisms*. Amsterdam: North-Holland Publishing Company, 1976.

Forster, K.I., and Chambers, S.M. Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, 1973, *12*, 627-635.

Fronek, J. Some criticisms of Halliday's 'information system.' *Lingua*, 1983, *60*, 311-329.

Givón, T. Topic, pronoun, and grammatical agreement. In C.N. Li (Ed.), *Subject and topic*. New York: Academic Press, 1976.

Givón, T. The drift from VSO to SVO in Biblical Hebrew: The pragmatics of tense-aspect. In C.N. Li (Ed.), *Mechanisms of syntactic change*. Austin: University of Texas, 1977.

Givón, T. From discourse to syntax: Grammar as a processing strategy. In T. Givón (Ed.), *Syntax and semantics* (Vol. 12). New York: Academic Press, 1979. (a)

Givón, T. *On understanding grammar*. New York: Academic Press, 1979. (b)

Givón, T. Topic continuity in discourse: An introduction. In T. Givón, *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam: John

Benjamins Publishing Company, 1983.

Glass, A.L., Holyoak, K.J., and Santa, J.L. *Cognition*. Reading, Massachusetts: Addison-Wesley Publishing Company, 1979.

Greenberg, J.H. Some universals of grammar with particular reference to the order of meaningful elements. In J.H. Greenberg (Ed.), *Universals of language* (2nd, ed.). Cambridge, Massachusetts: The M.I.T. Press, 1966.

Gregg, V.H. Word frequency, recognition and recall. In J. Brown (Ed.), *Recall and recognition*. New York: Wiley, 1976.

Gregg, V.H., Montgomery, D.C., and Castano, D. Recall of common and uncommon words from pure and mixed lists. *Journal of Verbal Learning and Verbal Behavior*, 1980, *19*, 240-245.

Grice, H.P. Logic and conversation. In P. Cole and J.L. Morgan (Eds.), *Syntax and semantics* (Vol. 3). New York: Academic Press, 1975.

Gundel, J.K. Role of topic and comment in linguistic theory. Bloomington, Indiana: Indiana University Linguistics Club, 1977.

Hall, J.F. Recognition as a function of word frequency. *American Journal of Psychology*, 1977, *92*, 497-505.

Halliday, M.A.K. Notes on transitivity and theme in English. *Journal of Linguistics*, 1967, *3*, 177-274.

Halliday, M.A.K. Language structure and language function. In J. Lyons (Ed.), *New horizons in linguistics*. Baltimore: Penguin Books, 1970.

Halliday, M.A.K. Modes of meaning and modes of expression: Types of grammatical structure and their determination by different semantic functions. In D.J. Anderson, E. Carney, and D. Holdcroft (Eds.), *Functions and context in linguistic analysis*. Cambridge: Cambridge University Press, 1979.

Hampton, J.A. An investigation of nature of abstract concepts. *Memory and Cognition*, 1981, *9*, 149-156.

Harris, M. Noun animacy and the passive voice: A developmental approach. *Quarterly Journal of Experimental Psychology*, 1978, *30*, 495-504.

Haviland, S.E., and Clark, H.H. What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1979, *13*, 512-521.

Hockett, C.F. *A course in modern linguistics.* New York: Macmillan, 1958.

Holland, V.M., and Rohrman, N.L. Distribution of the feature [+Animate] in the lexicon of the child. *Journal of Psycholinguistic Research*, 1979, *8*, 367-378.

Hornby, D.A. Surface structure and presupposition. *Journal of Verbal Learning and Verbal Behavior*, 1974, *13*, 530-538.

Hörmann, H. *To mean - To understand.* New York: Springer-Verlag, 1981.

James, C.T. The theme and imagery in the recall of active and passive sentences. *Journal of Verbal Learning and Verbal Behavior*, 1972, *11*, 205-211.

James, C.T., and Abrahamson, A.A. Recognition memory for active and passive sentences. *Journal of Psycholinguistic Research*, 1977, *6*, 37-47.

James, C.T., Thompson, J.G., and Baldwin, J.H. The reconstructive process in sentence memory. *Journal of Verbal Learning and Verbal Behavior*, 1973, *12*, 51-63.

Johnson, D.E. On Keenan's definition of "subject of." *Linguistic Inquiry*, 1977, *8*, 673-692. (a)

Johnson, D.E. On relational constraints on grammars. In P. Cole and J.M. Sadock (Eds.), *Syntax and semantics* (Vol. 8). New York: Academic Press, 1977. (b)

Johnson-Laird, P.N. The choice of the passive voice in communicative task. *British Journal of Psychology*, 1968, *59*, 7-15. (a)

Johnson-Laird, P.N. The interpretation of the passive voice. *Quarterly Journal of Experimental Psychology*, 1968, *20*, 69-73. (b)

Johnson-Laird, P.N. The passive paradox: A reply to Costermans and Huppet. *British Journal of Psychology*, 1977, *68*, 113-116.

Keenan, E.L. Towards a universal definition of subject. In C.N. Li (Ed.), *Subject and topic*, New York: Academic Press, 1976.

Keenan, E.L. The syntax of subject-final languages. In W.P. Lehmann (Ed.), *Syntactic typology*. Austin:

University of Texas Press, 1978.

Kempen, G. Conceptualizing and formulating in sentence production. In S. Rosenberg (Ed.), *Sentence production: Developments in research and theory*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977.

Kempen, G., and Huijbers, P. The lexicalization process in sentence production and naming: Indirect election of words. *Cognition*, 1983, *4*, 185-209.

Kieras, D. Beyond pictures and words: Alternative information - processing models of imagery effects on verbal memory. *Psychological Bulletin*, 1978, *85*, 532-554.

Kuno, S. Subject, theme, and the speaker's empathy: A reexamination of relativization phenomena. In C.N. Li (Ed.), *Subject and topic*. New York: Academic Press, 1976.

Kuno, S., and Kaburaki, E. Empathy and syntax. *Linguistic Inquiry*, 1977, *8*, 626-672.

Krupa, V. Syntactic typology and linearization. *Language*, 1982, *58*, 639-645.

Itagaki, N. *Contextual determinants of linguistic expressions: Point of view and given-new contraints on English*. Unpublished M.Sc. Thesis, University of Alberta, 1982.

Itagaki, N., and Prideaux, G.D. Pragmatic constraints on subject and agent selection. In S.C. Dik (Ed.), *Advances in functional grammar*. Dordrecht-Holland: Foris Publications, 1983.

Lachman, R. Uncertainty effects on time to access the internal lexicon. *Journal of Experimental Psychology*, 1973, *99*, 199-208.

Lachman, R., Lachman, J.M., and Butterfield, E.C. *Cognitive psychology and information processing: An introduction*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1979.

Labov, W. The boundaries of words and their meanings. In C.N. Bailey and R.W. Shuy (Eds.), *New ways of analyzing variation in English*. Washington, D.C.: Georgetown University Press, 1973.

Lakoff, G. Hedges: A study in meaning criteria and logic of fuzzy concepts. *Papers from the eighth regional* *meeting*. Chicago: Chicago Linguistics Society, 1972.

Levelt, W., and Massen, B. Lexical search and order of mention in sentence production. In W. Klein and W. Levelt (Eds.), *Crossing the boundaries in linguistics*. London: D. Reidel Publishing Company, 1981.

Li, C.N., and Thompson, S.A. Subject and topic: A new typology of language. In C.N. Li (Ed.), *Subject and topic*. New York: Academic Press, 1976.

Lyons, J. *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press, 1968.

Lyons, J. *Semantics* (Vol. 2). Cambridge: Cambridge University Press, 1977.

MacWhinney, B. Starting points. *Language*, 1977, *53*, 152-168.

MacWhinney, B., and Bates, E. Sentential devices for conveying givenness and newness: A cross cultural developmental study. *Journal of Verbal Learning and Verbal Behavior*, 1978, *17*, 539-558.

Malkiel, Y. *Essays on linguistic themes*. Oxford: Basil Blackwell, 1968.

Mathesius, V. On linguistic characterology. In J. Vachek (Ed.), *A Prague school reader in linguistics*. Bloomington, Indiana: Indiana University Press, 1964. (Originally published, 1928.)

Mathews, P.H. *Syntax*. Cambridge: Cambridge University Press, 1981.

Miller, G.A., and Johnson-Laird, P.N. *Language and perception*. Cambridge, Massachusetts: Harvard University Press, 1977.

Moeser, S.D. Memory for meaning and wording in concrete and abstract sentences. *Journal of Verbal Learning and Verbal Behavior*, 1974, *13*, 682-697.

Morley, G.D. Agentivity: A componential approach. *Lingua*, 1982, *60*, 177-182.

Neisser, U. *Cognitive Psychology*. Englewood Cliffs, New Jersey: Prentice-Hall, 1967.

Noble, C.E. An analysis of meaning. *Psychological Review*, 1952, *59*, 421-430.

Norman, D.A., and Rumelhart, D.E. *Explorations in cognition* San Francisco: W.H. Freeman and Company, 1975.

Ogden, C.K., and Richards, I.A. *The meaning of meaning.* London: Kegan Paul, 1923.

Osgood, C.E. Where do sentences come from? In D.A. Steinberg and L.A. Jakobovits (Eds.), *Semantics*: An interdisciplinary reader in philosophy, linguistics, and psychology. Cambridge: Cambridge University Press, 1971.

Osgood, C.E. *Lectures on language performance.* New York: Springer-Verlag, 1980.

Osgood, C.E., and Bock, J.K. Salience and sentencing: Some production principles. In S. Rosenberg (Ed.), *Sentence production: Developments in research and theory.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977.

Oldfield, R.C. Individual vocabulary and semantic currency: A preliminary study. *British Journal of Social and Clinical Psychology*, 1963, *2*, 122-130.

Oldfield, R.C. Things, words and the brain. *Quarterly Journal of Experimental Psychology*, 1966, *18*, 340-353.

Oldfield, R.C., and Wingfield, A. Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, 1965, *17*, 273-281.

Paivio, A. Mental imagery in associative learning and memory. *Psychological Review*, 1969, *76*, 241-263.

Paivio, A. *Imagery and verbal processes.* New York: Holt, Rinehart, and Winston, 1971.

Paivio, A., and Begg, I. *Psychology of language.* Englewood Cliffs, New Jersey: Prentice-Hall, 1981.

Paivio, A., Yuille, J.C., and Rogers, T.B. Noun imagery and meaningfulness in free and serial recall. *Journal of Experimental Pyschology*, 1967, *79*, 509-514.

Piaget, J. *The language and thought of the child.* New York: Harcourt Brace Jovanovich, 1926.

Piaget, J. *The child's conception of the world.* New York: Harcourt Brace Jovanovich, 1929.

Piaget, J. Piaget's theory. In P.H. Mussen (Ed.), *Carmichael's manual of child psychology* (Vol. 1). New York: John Wiley & Sons, 1970.

Porzig, W. Wessenhafte beheutungsbeziehungen. *Beiträge zur Geschichte der deutsuchen Sprache und Literature*, 1934, *58*, 70-97.

Pullum, G.K. Word order universal and grammatical relations. In P. Cole and J.M. Sadock (Eds.), *Syntax and Semantics* (Vol. 8). New York: Academic Press, 1977.

Quillian, M.R. Semantic memory. In M. Minsky (Ed.), *Semantic information processing*. Cambridge: Massachusatts: The M.I.T. Press, 1968.

Quirk, R., and Greenbaum, S. *A university grammar of English*. London: Longman, 1973.

Ransom, E.N. Definiteness, animacy, and NP ordering. *Proceedings of the Third Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society, 1977.

Rechards, L.G. Concreteness as a variable in word recognition. *American Journal of Psychology*, 1976, *89*, 707-718.

Reynolds, A., and Pavio, A. Cognitive and emotional determinants of speech. *Canadian Journal of Psychology*, 1968, *22*, 164-175.

Rohrman, N.L. More on the recall of nominalization. *Journal of Verbal Learning and Verbal Behavior*, 1970, *9*, 534-536.

Rosch, E. On the internal structure of perceptual and semantic categories. In T.E. Moore (Ed.), *Cognitive development and the acquisition of language*. New York: Academic Press, 1973.

Rosch, E. Linguistic relativity. In A. Silverstein (Ed.), *Human communication: Theoretical perspective*, New York: Halsted Press, 1974.

Rosch, E. Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 1975, *104*, 192-233.

Rosch, E. Principles of categorization. In E. Rosch and B.B. Lloyd (Eds.), *Cognition and categorization*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1978.

Rosch, E., Mervis, C.B., Gary, W., Johnson, D., and Boyes-Braem, P. Basic objects in natural categories. *Cognitive Psychology*, 1976, *8*, 382-439.

Russell, R.W. Studies in animism: II. The development of animism. *The Journal of Genetic Psychology*, 1940, *56*, 353-366. (a)

Russell, R.W. Studies in animism: IV. An investigation of concepts allied to animism. *The Journal of Genetic Psychology*, 1940, *57*, 83-91. (b)

Russell, R.W. Studies in animism: V. Animism in older children. *The Journal of Genetic Psychology*, 1942, *60*, 329-335.

Russell, R.W., and Dennis, W. Studies in animism: I. A standardized procedure for the investigation of animism. *The Journal of Genetic Psychology*, 1939, *55*, 389-400.

Sandmann, M. *Subject and predicate: A contribution to the theory of syntax* (2nd rev. ed.). Heidelberg: CarlWinter Universitatsverlag, 1979.

Saussure, F. de *A course in general linguistics*. New York: McGraw-Hill, 1916.

Schank, R.C. Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, 1972, *3*, 552-631.

Schmidt, C. The relevance to semantic theory of a study of vagueness. *Papers from the tenth regional meeting*. Chicago: Chicago Linguistics Society, 1974.

Smith, E.E., and Medin, D.L. *Categories and concepts*. Cambridge, Massachusetts: Harvard University Press, 1981.

Steele, S. Word order variation: A typological study. In J.H. Greenberg (Ed.), *Universals of human language* (Vol. 4), Stanford: Stanford University Press, 1978.

Talmy, L. Figure and ground in complex sentence. In J. H. Greenberg (Ed.), *Universals of human language* (Vol. 4), Stanford: Stanford University Press, 1978.

Thorndike, E.L., and Lorge, I. *The teacher's word book of 30,000 words*. New York: Bureau of Publications, Teacher's College, 1944.

Tomlin R.S. *An explanation of the distribution of basic constituent orders*. Unpublished Doctoral Dissertation, University of Michigan, 1979.

Tomlin, R.S. On the interaction of syntactic subject, thematic information, and agent in English. *Journal of*

Pragmatics, 1983, *7*, 411-432.

Tulving, E. Episodic and semantic memory. In E. Tulving and W. Donaldson (Eds.), *Organization of memory*, New York: Academic Press, 1972.

Ultan, R. Some general characteristics of interrogative systems. *Working Papers on Language Universals*, 1969, *1*, 41-63.

Underwood, B.J., and Schulz, R.W. *Meaningfulness and verbal learning*. Chicago: J.B. Lippincott Company, 1960.

Vachek, J. *The linguistic school of Prague*. Bloomington, Indiana: Indiana University Press, 1966.

Vincent, N. Word order and grammatical theory. In J.M. Meisel and M.D. Pam (Eds.), *Linear order and generative theory*. Amsterdam: John Benjamins B.V., 1979.

Yuillie, J.C., and Paivio, A. Latency of imagery and verbal mediators as a function of stimulus and response concreteness - imagery. *Journal of Experimental Psychology*, 1967, *75*, 540-544.

Walter D.A., and Fox, D. The role of perceptual and semantic elaboration in high- and low-imagery sentence recall. *Memory and Cognition*, 1981, *9*, 625-630.

Weiner, E.J., and Labov, W. Constraints on the agentless passive. *Journal of Linguistics*, 1983, *19*, 29-58.

Zipf, G.K. *The psycho-biology of language*. Boston: Houghton Mifflin, 1935.

Zubin, D.A. Discourse function of morphology: The focus system in German. In T. Givón (Ed.), *Syntax and semantics* (Vol. 12). New York: Academic Press, 1979.

APPENDIX A: Nouns Used in Experiment 1

The numbers in the second and third columns below indicate the frequency counts from Carroll, Davies, and Richman (1971), i.e., the numbers of the times the word types occurred in the total corpus of 5,088,727 tokens, and from Thorndik and Lorg (1944), i.e., the numbers of the times the word types occurred per million words ("A" and "AA" indicate "at leat 50 per million words" and "100 or over per million words").

HHUM			LHUM		
boy	2941	AA	buffoon	2	1
professor	154	A	foundling	0	0
woman	1492	AA	mentor	0	1
driver	481	40	outcat	6	3
player	490	36	sniper	2	0
student	789	A	highwayman	10	1
worker	815	A	idler	1	2
writer	451	A	yardman	0	0
captain	815	AA	anarchist	2	2
doctor	795	AA	charlatan	4	1

HANH			LANH		
dog	2161	AA	wombat	3	0
horse	2161	AA	gnu	1	0
sheep	471	A	jackal	0	2
cat	943	A	nighthawk	2	0
bear	937	AA	tsetsefly	0	0
cow	598	A	albatross	12	1
robin	345	48	dodo	0	0
ant	482	38	piranha	6	0
lion	464	A	weevil	13	2
bee	335	A	haddock	6	1

To be continued

HOBJ			LOBJ		
car	2524	AA	ricksha(w)	0	0
book	2428	AA	hydroplane	1	0
table	1737	AA	plutonium	4	0
ship	1815	AA	hutch	5	0
gold	929	AA	wampum	13	1
river	2468	AA	papaya	0	0
coat	512	AA	peashooter	1	0
apple	582	A	sycamore	6	4
gun	605	A	mandolin	2	0
rose	520	AA	kumquat	0	0

HABS			LABS		
fact	1450	AA	affluence	3	1
idea	1399	AA	acquiescence	2	2
attention	476	AA	dejection	1	2
freedom	397	A	rapprochement	2	1
peace	385	AA	rationalization	0	0
future	364	AA	innuendo	0	1
spirit	351	AA	technocracy	0	2
education	320	A	dominance	3	1
religion	221	A	virtuosity	9	0
knowledge	424	AA	comradeship	2	1

APPENDIX B: Sample Single Sentences

HHUM: "drivers," "players," and "students."

a. *I met the bus driver.*
b. *The student did not agree with the taxi driver regarding the fare.*

a. *He is a terrible guiter player.*
b. *The hockey player was known to most of the country for his skating ability.*

a. *The student went to the library.*
b. *The students of the university elected a new president yesterday.*

LHUM: "snipers," "highwaymen," and "idlers."

a. *The well hidden sniper fired four shots at the oncoming truck.*
b. *A sniper killed J. F. Kennedy.*

a. *The highwayman went out of business when motor cars came along.*
b. *To hold up stage couchs and rob from the rich was the employment of a highwayman.*

a. *The idler was seen in the park.*
b. *An idler doesn't go very far in university unless he has professors who are also lazy.*

HANH: "cats," "bears," and "cows."

a. *The fawn-coloured Persian cat in the pet shop window was not the right pet for the child.*
b. *When I am exposed to a cat, I have an allergic reaction.*

a. *The black bear of North America is not often seen in the winter months because of the habit of hibernation.*
b. *Where I live, a bear is a common sight and a fact of life.*

a. *The jersey cow is perhaps the most docile of all domesticated cows.*
b. *The other day, my son and I saw a caribou cow in the wild.*

LANH: "piranhas," "tsetseflies," and "albatrosses."

a. *I would not consider swimming in a river where there are men eating fish, such as piranhas.*
b. *The piranha is a dangerous fish in groups.*

a. *The tsetsefly bit the explorer.*
b. *The tsetsefly is a well-known cause of narcolepsy.*

a. *An albatross followed the boat.*
b. *I have never seen an albatross at sea or anywhere else.*

HOBJ: "tables," "coats," and "apples."

a. *I am using my coffee table for a footstool while doing this task.*
b. *The highly ornate table was sold at the auction last Saturday for twenty-five dollars.*

a. *He put on his coat.*
b. *The expensive fur coat was hanging by itself on a hang in the closet.*

a. *That tree has the best apples.*
b. *An apple a day keeps the doctor away, Ha!*

LOBJ: "hutches," "wampum," and "papayas."

a. *The rabbits were quite content in their hutch.*
b. *He fogot to close the rabbit hatch.*

a. *The Venetian's currency was easier to mint than the wampum of the North American Indian.*
b. *The Indian paid wampum for the guns.*

a. *Papaya fruit is not one of my favorites.*
b. *He ate the papaya for breakfast.*

HABS: "attention," "freedom," and "peace."

a. *It is very hard to get the attention of a two-year child.*
b. *"Let me draw your attention to the following facts," said the professor.*

a. *Freedom is what democracy is all about.*
b. *In 1861, freedom was a personal liberty that was rare in the confederate state of America.*

a. *It seems to me that as long as there are humans, there will never be peace on earth.*

b. *A peace settlement in the Middle East would be possible if all foreign troops left.*

LABS: "rationalization," "innuendo," and "technocracy."

a. *The rationalization for this arguement is quite simple.*
b. *Rationalization is a psychological pehnomenon.*

a. *She hated the innuendo of it all.*
b. *Innuendo is the coward's way of saying what he really feels.*

a. *He didn't agree with the theory of technocracy.*
b. *General Motors is run by a techocracy.*

APPENDIX C: Sample Passages

HHUM: "women" and "boys."

The woman was patiently waiting for the bus. Too bad no one told her the bus wasn't running that day.

Boys will be boys. This is a rather dangerous statement to make in the 1980's. What used to be a safe definition of a boy is now contested by feminists, machos and even homosexuals.

LHUM: "outcasts" and "buffoons."

He was used to being an outcast, ever since he started school, it seems, he hadn't fit in. Now he seemed to be boxed into the role. He's so used to be the outcat he came to expect it.

Gerald Ford, former president, was often described in the unkindest terms. One of the worst insults was when he appeared on the cover of a magazine dressed as a buffoon.

HANH: "bears" and "dogs."

I wonder if there's a difference between Eastern and Western bears. Here they seem to be so feared. In the East they were sort of cute.

Man's best friend is his dog. Most people will admit that dogs are more trustworthy than cats, yet cats seem to imspire most creative activity.

LANH: "gnus" and "dodos."

The gnu and the antelope met one day in the forest. Neither knew which one was in the wrong area, in fact neither knew how they got there.

Who would believe that man would go the way of the dodo bird? With his ever-growing lust for control over nature, man is on the verge of destroying not only his environment but in consequence his own playground.

HOBJ: "books" and "guns."

We have so many unread books on our shelves, it's incredible. I guess it comes from two definitions of a library. To me a home library is composed of books which have been read and enjoyed. To my husband it is composed of many books so there is an adequate selection.

Guns always evoke images of western movies. What a waste of time! I am gun shy and can't see what people see in guns. But our neighbours to the South love them.

LOBJ: "peashooters" and "plutonium."

The southern U.S., in early 1900's, is about the only place I can contextualize peashooters. I think of boys hiding arounds a big house shooting at each other with peashooters.

Well do I remember the sage words of a fellow student in Grade 11: nucleur scientists should be all buried together in a big hole. And what do you think about the discoverer of plutonium?

HABS: "fact" and "education."

From watching police and law type shows it seems a fact's real value depends so much on its context. I hope it really isn't as valuable as it seems on TV and in movies.

Education is mandatory in Western democracies but few rarely speak about learning. As soon as one gets his education, all learning seems to cease. I learned this at school.

LABS: "affluence" and "innuendo."

Alberta is such a province of affluent people. Even in these recessionary times the affluence is still so evident.

Modern writers always use innuendo without intending to do so. They want to imply certain things but are forever leaving out things, hoping that we poor readers will.

APPENDIX D: Mean Ratios for Nouns

HHUM	Single	Passage	LHUM	Single	Passage
man	0.618	0.671	buffoon	0.344	0.296
professor	0.588	0.736	founding	0.422	0.500
woman	0.781	0.761	mentor	0.313	0.359
driver	0.627	0.833	outcast	0.294	0.328
player	0.719	0.654	sniper	0.563	0.792
student	0.647	0.455	highwayman	0.700	0.626
worker	0.619	0.724	idler	0.441	0.474
writer	—	0.400	yardman	0.647	0.718
captain	0.729	0.679	anarchist	0.219	0.292
doctor	0.333	0.574	charlatan	0.400	0.485

HANH	Single	Passage	LANH	Single	Passage
dog	0.696	0.840	wombat	0.563	0.546
horse	0.500	0.472	gnu	0.611	0.611
sheep	0.500	0.477	jackal	0.676	0.933
cat	0.734	0.610	nighthawk	0.778	0.476
bear	0.507	0.652	tsetsefly	0.559	0.615
cow	0.676	0.514	albatross	0.750	0.671
robin	0.765	0.703	dodo	0.824	0.572
ant	0.679	0.442	piranha	0.813	0.538
lion	0.615	0.603	weevil	0.656	0.464
bee	0.533	0.814	haddock	0.367	0.524

(To be continued)

HOBJ	Single	Passage	LOBJ	Single	Passage
car	0.656	0.455	ricksha(w)	0.313	0.231
book	0.375	0.506	hydroplane	0.469	0.680
table	0.533	0.444	plutonium	0.647	0.394
ship	0.333	0.571	hutch	0.250	0.288
gold	0.412	0.354	wampum	0.219	0.410
river	0.375	0.646	papaya	0.471	0.468
coat	0.353	0.400	peashooter	0.188	0.310
apple	0.465	0.455	sycamore	0.333	0.458
gun	0.333	0.548	mandolin	0.469	0.299
rose	0.750	0.551	kumquat	0.531	0.455

HABS	Single	Passage	LABS	Single	Passage
fact	0.375	0.461	affluence	0.375	0.681
idea	0.353	0.436	acquiescence	0.667	0.361
attention	0.063	0.107	dejection	0.353	0.551
freedom	0.667	0.381	rapprochement	0.467	0.591
peace	0.633	0.321	rationalization	0.563	0.361
future	0.265	0.564	innuendo	0.400	0.372
spirit	0.571	0.409	technocracy	0.433	0.577
education	0.375	0.429	dominance	0.625	0.191
religion	0.656	0.643	virtuosity	0.500	0.327
knowledge	0.176	0.308	comradeship	0.500	0.393

APPENDIX E: Detailed Results of *A Posteriori* Tests

The first and second tables below present the studentized ranges, Q values, obtained on the basis of the ratio and percentage measurements, respectively. (The values below the labels of the word groups indicate the mean transformed ratio or percentage measurements for the corresponding word groups.)

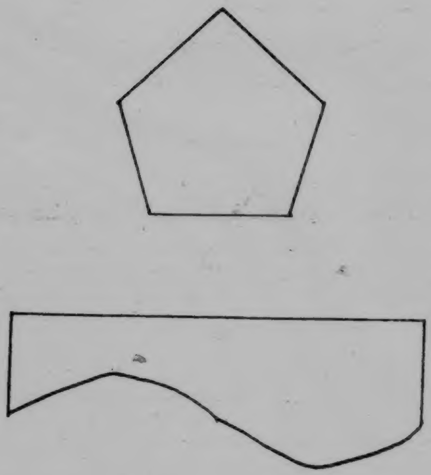
	LOBJ 1.13	HABS 1.15	LABS 1.22	LHUM 1.23	HOBJ 1.25	HANH 1.44	HHUM 1.49	LANH 1.51
LOBJ	—	0.40	1.93	2.13	2.53	6.94***	8.03***	8.43***
HABS		—	1.53	1.73	2.13	6.54***	7.63***	8.03***
LABS			—	0.20	0.60	5.01***	6.10***	6.50***
LHUM				—	0.40	4.81***	5.90***	6.30***
HOBJ					—	4.41***	5.50***	5.90***
HANH						—	1.09	1.49
HHUM							—	0.40
LANH								—

	HABS 0.91	LOBJ 0.95	LABS 0.97	LHUM 1.04	HOBJ 1.06	LANH 1.28	HANH 1.29	HHUM 1.31
HABS	—	0.96	1.56	3.40*	3.89*	9.58***	9.89***	10.5***
LOBJ		—	0.60	2.44	2.93	8.62***	8.93***	9.53***
LABS			—	1.84	2.34	8.02***	8.33***	8.93***
LHUM				—	0.49	6.18***	6.49***	7.09***
HOBJ					—	5.69***	6.00***	6.59***
LANH						—	0.31	0.91
HANH							—	0.60
HHUM								—

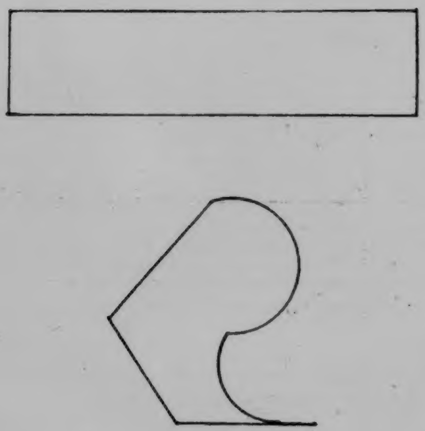
*$p<0.10$
***$p<0.01$

APPENDIX F: Pictures Used in Experiments 2 and 3

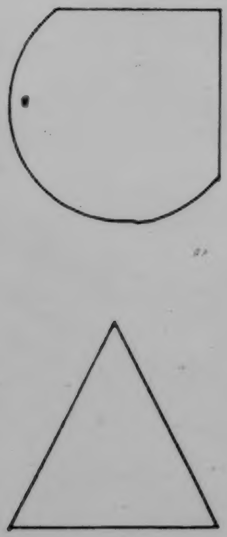
UV1



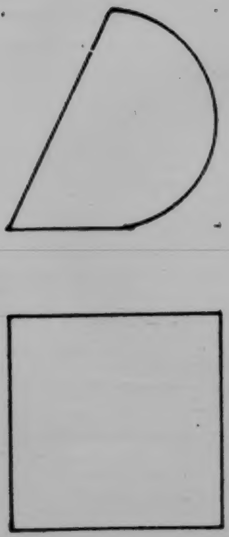
UV2



MV1

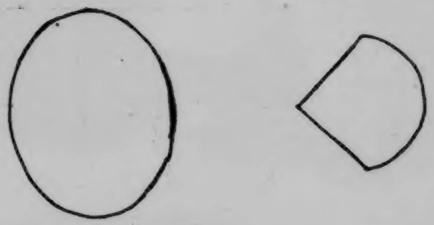


MV2

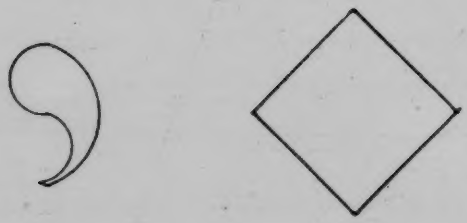


(To be continued)

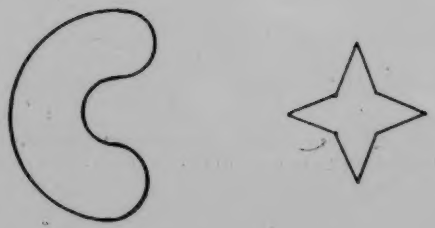
US1



US2



MS1



MS2